

Distortion Bounds for Source Broadcast Problem

Lei Yu, Houqiang Li, *Senior Member, IEEE*, and Weiping Li, *Fellow, IEEE*

Abstract

This paper investigates the joint source-channel coding problem of sending a memoryless source over a memoryless broadcast channel. An inner bound and several outer bounds on the achievable distortion region are derived, which respectively generalize and unify several existing bounds. As a consequence, we also obtain an inner bound and an outer bound for degraded broadcast channel case. When specialized to Gaussian source broadcast or binary source broadcast, the inner bound and outer bound not only recover the best known inner bound and outer bound in the literature, but also are used to generate some new results. Besides, we also extend the inner bound and outer bounds to Wyner-Ziv source broadcast problem, i.e., source broadcast with side information available at decoders. Some new bounds are obtained when specialized to Wyner-Ziv Gaussian case and Wyner-Ziv binary case. In addition, when specialized to lossless transmission of a source with independent components, the bounds for source broadcast problem (without side information) are also used to achieve an inner bound and an outer bound on capacity region of general broadcast channel with common messages, which respectively generalize Marton's inner bound and Nair-El Gamal outer bound to K -user broadcast channel case.

Index Terms

Joint source-channel coding (JSCC), hybrid coding, hybrid digital-analog (HDA), broadcast, Wyner-Ziv, side information, multivariate covering/packing, network information theory.

I. INTRODUCTION

As stated in Shannon's source-channel separation theorem [2], cascading source coding and channel coding does not lose the optimality for the point-to-point communication systems. This separation theorem does not only suggest a simple system architecture in which source coding and channel coding are separated by a universal digital interface, but also guarantees that such architecture does not incur any asymptotic performance loss. Consequently, it forms the basis of the architecture of today's communication systems. However, for many multi-user communication systems, the optimality of such a separation does not hold any more [3], [4]. Therefore, an increasing amount of literature focus on joint source-channel coding (JSCC) in multi-user setting.

The authors are all with the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027, China (e-mail: yulei@ustc.edu.cn, lihq@ustc.edu.cn, wpli@ustc.edu.cn). The material in this paper was presented in part at IEEE ISIT 2016 [1].

One of the most classical problems in this area is JSCC of transmitting a Gaussian source over average power constrained K -user Gaussian broadcast channel. Gobblick [3] observed that when the source bandwidth and the channel bandwidth are matched (i.e., one channel use per source sample) linear uncoded transmission (symbol-by-symbol mapping) is optimal. However, the optimality of such a simple linear scheme cannot be extended to the case of the bandwidth mismatch. One way to approximately characterize the achievable distortion region is finding its inner bound and outer bound. For inner bound, analog coding schemes or hybrid coding schemes have been studied in a vast body of literature [4], [5], [6], [7]. For 2-user Gaussian broadcast communication, Prabhakaran *et al.* [7] gave the tightest inner bound so far, which is achieved by hybrid digital-analog (HDA) scheme. On the other hand, Reznic *et al.* [8] derived a nontrivial outer bound for 2-user Gaussian broadcast problem with bandwidth expansion (i.e., more than one channel uses per source sample) by introducing an auxiliary random variable (or remote source). Tian *et al.* [9] extended this outer bound to K -user case by introducing more than one auxiliary random variables. Similar to the results of Reznic *et al.*, the outer bound given by Tian *et al.* is also nontrivial only for bandwidth expansion case [10]. Beyond broadcast communication, Minero *et al.* [15] considered sending memoryless correlated source transmitted over memoryless multi-access channel and derived an inner bound using a unified framework of hybrid coding, and also Lee *et al.* [20] derived a unified achievability result for memoryless network communication.

Besides, in [6], [17], [18] Wyner-Ziv source communication problem was investigated, in which side information of the source is available at decoder(s). Shamai *et al.* [6] studied the problem of sending Wyner-Ziv source over point-to-point channel, and proved that for such communication system separate coding (which combines Wyner-Ziv coding with channel coding) does not incur any loss of optimality. Nayak *et al.* [17] and Gao *et al.* [18] investigated Wyner-Ziv source broadcast problem, and obtained an outer bound by simply applying cut-set bound (the minimum distortion achieved in point-to-point setting) for each receiver.

In this paper, we consider JSCC of transmitting a memoryless source over K -user memoryless broadcast channel, and give an inner bound and three outer bounds on the achievable distortion region. The inner bound is derived by a unified framework of hybrid coding inspired by [15], and the outer bounds are derived by introducing auxiliary random variables at sender side or at receiver sides. The proof method of introducing auxiliary random variables (or remote sources) at sender side could be found in [8] and [9]. However, to the best of our knowledge, it is the first time to prove outer bounds by introducing auxiliary random variables (or remote channels) at receiver sides. Our bounds are generalizations and unifications of several existing bounds in the literature. Besides, as a consequence, we also obtain an inner bound and an outer bound for degraded broadcast channel case. Owing to the generalization of our results, when specialized to Gaussian source broadcast and binary source broadcast, our inner bound could recover best known performance achieved by hybrid coding, and our outer bound could recover the best known outer bounds given by Tian *et al.* [9] and Khezeli *et al.* [11]. Moreover, for these cases, our bounds can also be used to generate some new results. Besides, we also extend the inner bound and outer bounds to Wyner-Ziv source broadcast problem, i.e., source broadcast with side information at decoders. When specialized to Wyner-Ziv Gaussian case and Wyner-Ziv binary case, our bounds reduce to some new bounds. In addition, when

specialized to lossless transmission of a source with independent components, the bounds for source broadcast problem (without side information) is also used to achieve an inner bound and an outer bound on capacity region of general broadcast channel with common messages, which respectively generalize Marton's inner bound and Nair-El Gamal outer bound to K -user broadcast channel case.

The rest of this paper is organized as follows. Section II summarizes basic notations, definitions and preliminaries, and formulates the problem. Section III gives the main results for source broadcast problem, including general, degraded, Gaussian and binary cases. Section IV extends the results to Wyner-Ziv source broadcast problem. Finally, Section V gives the concluding remarks.

II. PROBLEM FORMULATION AND PRELIMINARIES

A. Notation

Throughout this paper, we follow the notation in [16]. For example, for discrete random variable $X \sim p_X$ on alphabet \mathcal{X} and $\epsilon \in (0, 1)$, the set of ϵ -typical n -sequences x^n (or the typical set in short) is defined as $\mathcal{T}_\epsilon^{(n)}(X) = \{x^n : |\{i : x_i = x\}|/n - p_X(x)| \leq \epsilon p_X(x) \text{ for all } x \in \mathcal{X}\}$. When it is clear from the context, we will use $\mathcal{T}_\epsilon^{(n)}$ instead of $\mathcal{T}_\epsilon^{(n)}(X)$.

In addition, we use $X_{\mathcal{A}}$ to denote the vector $(X_j : j \in \mathcal{A})$, use $[i : j]$ to denote the set $\{[i], [i] + 1, \dots, [j]\}$, and use $\mathbf{1}$ to denote an all-one vector (similarly, use $\mathbf{2}$ to denote an all-2 vector). We say vector $m_{[1:N]}$ is smaller than vector $m'_{[1:N]}$ if $m_j = m'_j, k < j \leq K$ and $m_k < m'_k$ for some k . For two vectors $m_{\mathcal{I}}$ and $m'_{\mathcal{I}}$, we say $m_{\mathcal{I}}$ is component-wise unequal to $m'_{\mathcal{I}}$, if $m_i \neq m'_i$ for all $i \in \mathcal{I}$, and denote it as $m_{\mathcal{I}} \not\equiv m'_{\mathcal{I}}$. Besides, we use $1\{\mathcal{A}\}$ to denote indicator function of event \mathcal{A} , i.e.,

$$1\{\mathcal{A}\} = \begin{cases} 1, & \text{if } \mathcal{A} \text{ is true;} \\ 0, & \text{if } \mathcal{A} \text{ is false.} \end{cases}$$

B. Problem Formulation

Consider the source broadcast system shown in Fig. 1. The discrete memoryless source (DMS) S^n is first coded into X^n using a source-channel code, then transmitted to K receivers through a discrete memoryless broadcast channel (DM-BC) $p_{Y_{[1:K]}|X}$, and finally, the receiver k produces source reconstruction \hat{S}_k^n from the received signal Y_k^n .

Definition 1 (Source). A discrete memoryless source (DMS) is specified by a probability mass function (pmf) p_S on a finite alphabet \mathcal{S} . The DMS p_S generates an i.i.d. random process $\{S_i\}$ with $S_i \sim p_S$.

Definition 2 (Broadcast Channel). A K -user discrete memoryless broadcast channel (DM-BC) is specified by a collection of conditional pmfs $p_{Y_{[1:K]}|X}$ on finite output alphabet $\mathcal{Y}_1 \times \dots \times \mathcal{Y}_K$ for each x in finite input alphabet \mathcal{X} .

Definition 3 (Degraded Broadcast Channel). A DM-BC $p_{Y_{[1:K]}|X}$ is stochastically degraded (or simply degraded) if there exist a random vector $\tilde{Y}_{[1:K]}$ such that $\tilde{Y}_k | \{X = x\} \sim p_{Y_k|X}(\tilde{y}_k|x), 1 \leq k \leq K$, i.e., $\tilde{Y}_{[1:K]}$ has the same

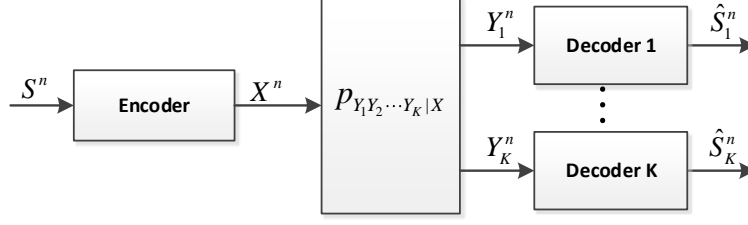


Fig. 1. Source broadcast system.

conditional marginal pmfs as $Y_{[1:K]}$ (given X), and $X \rightarrow \tilde{Y}_K \rightarrow \tilde{Y}_{K-1} \rightarrow \dots \rightarrow \tilde{Y}_1$ ¹ form a Markov chain. In addition, as a special case, if $X \rightarrow Y_K \rightarrow Y_{K-1} \rightarrow \dots \rightarrow Y_1$, i.e., $\tilde{Y}_k = Y_k, 1 \leq k \leq K$, then $p_{Y_{[1:K]}|X}$ is physically degraded.

Definition 4. An n -length source-channel code is defined by the encoding function $x^n : \mathcal{S}^n \mapsto \mathcal{X}^n$ and a sequence of decoding functions $\hat{s}_k : \mathcal{Y}_k^n \mapsto \hat{\mathcal{S}}_k^n, 1 \leq k \leq K$, where $\hat{\mathcal{S}}_k$ is the alphabet of source reconstruction at receiver k .

For any n -length source-channel code, the induced distortion is defined as

$$\mathbb{E}d_k(S^n, \hat{S}_k^n) = \frac{1}{n} \sum_{t=1}^n \mathbb{E}d_k(S_t, \hat{S}_{k,t}), \quad (1)$$

for $1 \leq k \leq K$, where $d_k(s, \hat{s}_k) : \mathcal{S} \times \hat{\mathcal{S}}_k \mapsto [0, +\infty]$ is a distortion measure function for receiver k .

Definition 5. For transmitting source S over channel $p_{Y_{[1:K]}|X}$, if there exists a sequence of source-channel codes such that

$$\limsup_{n \rightarrow \infty} \mathbb{E}d_k(S^n, \hat{S}_k^n) \leq D_k, \quad (2)$$

then we say that the distortion tuple $D_{[1:K]}$ is achievable.

Definition 6. For transmitting source S over channel $p_{Y_{[1:K]}|X}$, the admissible distortion region is defined as

$$\mathcal{R} \triangleq \{D_{[1:K]} : D_{[1:K]} \text{ is achievable}\}. \quad (3)$$

The admissible distortion region \mathcal{R} only depends on the marginal distributions of $p_{Y_{[1:K]}|X}$, hence for source broadcast over stochastically degraded channel it suffices to only consider the broadcast over physically degraded channel.

In addition, Shannon's source-channel separation theorem shows that the minimum distortion for transmitting source over point-to-point channel satisfies

$$R_k(D_k) = C_k, \quad (4)$$

¹To simplify notation, the Markov chain is assumed in this direction. Note that this differs from that in the conference version [1].

where $R_k(\cdot)$ is the rate-distortion function of the source with distortion measure d_k , and C_k is the capacity of the channel of the receiver k . Therefore, the optimal distortion (Shannon limit) is

$$D_k^* = R_k^{-1}(C_k). \quad (5)$$

Obviously,

$$\mathcal{R} \subseteq \mathcal{R}^* \triangleq \{D_{[1:K]} : D_k \geq D_k^*, 1 \leq k \leq K\}, \quad (6)$$

where \mathcal{R}^* is named *trivial outer bound*.

In the system above, source bandwidth and channel bandwidth are matched. In this paper, we also consider the communication system with bandwidth mismatch, whereby m samples of a DMS are transmitted through n uses of a DM-BC. For this case, bandwidth mismatch factor is defined as $b = \frac{n}{m}$.

C. Multivariate Covering/Packing Lemma

Two important results we need to prove the achievability part in this work are the following lemmas, both of which are generalized versions of the existing covering/packing lemmas.

Let $(U, V_{[0:k]}) \sim p_{U, V_{[0:k]}}$, and let $(U^n, V_0^n) \sim p_{U^n, V_0^n}$ be a random vector sequence. For each $j \in [1 : k]$, let $\mathcal{A}_j \subseteq [1 : j-1]$. Assume \mathcal{A}_j satisfies if $i \in \mathcal{A}_j$, then $\mathcal{A}_i \subseteq \mathcal{A}_j$. For each $j \in [1 : k]$ and each $m_{\mathcal{A}_j} \in \prod_{i \in \mathcal{A}_j} [1 : 2^{nr_i}]$, let $V_j^n(m_{\mathcal{A}_j}, m_j), m_j \in [1 : 2^{nr_j}]$, be pairwise conditionally independent random sequences, each distributed according to $\prod_{i=1}^n p_{V_j|V_{\mathcal{A}_j}, V_0}(v_{j,i}|v_{\mathcal{A}_j,i}(m_{\mathcal{A}_j}), v_{0,i})$. Hence for each $j \in [1 : k]$, $\mathcal{A}_j \cup \{0\}$ denotes the index set of the random variables on which the codeword V_j^n is superposed. Based on the notations above, we have the following generalized Multivariate Covering Lemma and generalized Multivariate Packing Lemma.

Lemma 1 (Multivariate Covering Lemma). *Let $\epsilon' < \epsilon$. If $\lim_{n \rightarrow \infty} \mathbb{P}\left((U^n, V_0^n) \in \mathcal{T}_{\epsilon'}^{(n)}\right) = 1$, then there exists $\delta(\epsilon)$ that tends to zero as $\epsilon \rightarrow 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left((U^n, V_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_{\epsilon}^{(n)} \text{ for some } m_{[1:k]}\right) = 1, \quad (7)$$

if $\sum_{j \in \mathcal{J}} r_j > \sum_{j \in \mathcal{J}} H(V_j|V_{\mathcal{A}_j} V_0) - H(V_{\mathcal{J}}|V_0 U) + \delta(\epsilon)$ for all $\mathcal{J} \subseteq [1 : k]$ such that $\mathcal{J} \neq \emptyset$ and if $j \in \mathcal{J}$, then $\mathcal{A}_j \subseteq \mathcal{J}$.

Lemma 2 (Multivariate Packing Lemma). *There exists $\delta(\epsilon)$ that tends to zero as $\epsilon \rightarrow 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left((U^n, V_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_{\epsilon}^{(n)} \text{ for some } m_{[1:k]}\right) = 0, \quad (8)$$

if $\sum_{j \in \mathcal{J}} r_j < \sum_{j \in \mathcal{J}} H(V_j|V_{\mathcal{A}_j} V_0) - H(V_{\mathcal{J}}|V_0 U) - \delta(\epsilon)$ for some $\mathcal{J} \subseteq [1 : k]$ such that $\mathcal{J} \neq \emptyset$ and if $j \in \mathcal{J}$, then $\mathcal{A}_j \subseteq \mathcal{J}$.

Note that all the existing covering and packing lemmas such as [16, Lem. 8.2] and [19, Lem. 4], only involve single-layer codebook. Our Multivariate Covering and Packing Lemmas generalize them to the case of multilayer codebook, and certainly our Covering/Packing Lemmas could recover all of them.

III. SOURCE BROADCAST

A. Source Broadcast

Now, we bound the distortion region for source broadcast communication. To write the inner bound, we first introduce an auxiliary random variable $V_j, 1 \leq j \leq N \triangleq 2^K - 1$ for each of the $2^K - 1$ nonempty subsets $\mathcal{G}_j \subseteq [1 : K]$, and let V_j denote a common message transmitted from sender to all the receivers in \mathcal{G}_j . The V_j corresponds to a subset \mathcal{G}_j by the following one-to-one mapping.

Sort all the nonempty subsets $\mathcal{G} \subseteq [1 : K]$ in the decreasing order². Map the j th subset in the resulting sequence to j . Obviously this mapping is one-to-one corresponding. For example, if $K = 3$, then $\mathcal{G}_1 = \{1, 2, 3\}, \mathcal{G}_2 = \{2, 3\}, \mathcal{G}_3 = \{1, 3\}, \mathcal{G}_4 = \{1, 2\}, \mathcal{G}_5 = \{3\}, \mathcal{G}_6 = \{2\}, \mathcal{G}_7 = \{1\}$.

Besides, let

$$\mathcal{A}_j \triangleq \{i \in [1 : N] : \mathcal{G}_j \subsetneq \mathcal{G}_i\}, 1 \leq j \leq N, \quad (9)$$

$$\mathcal{D}_k \triangleq \{i \in [1 : N] : k \in \mathcal{G}_i\}, 1 \leq k \leq K. \quad (10)$$

Later we will show that they respectively correspond to the index set of the random variables on which the codeword V_j^n is superposed, and the index set of decodable codewords V_j^n 's for receiver k in the proposed hybrid coding scheme; see Appendix C-A. Decoder k is able to recover correctly the V_j^n , designated by the encoder with probability approaching 1 as $n \rightarrow \infty$ if $j \in \mathcal{D}_k$. In addition, it is easy to verify that if $j \in \mathcal{D}_k$, then $\mathcal{A}_j \subseteq \mathcal{D}_k$. It means that the proposed codebook satisfies that if information V_j^n can be recovered correctly by receiver k , then $V_{\mathcal{A}_j}^n$ can also be recovered correctly by it.

Based on the notations above, we define distortion region (inner bound)

$$\begin{aligned} \mathcal{R}^{(i)} = & \left\{ D_{[1:K]} : \text{There exist some pmf } p_{V_{[1:N]}|S}, \text{ vector } r_{[1:N]}, \right. \\ & \text{and functions } x(v_{[1:N]}, s), \hat{s}_k(v_{\mathcal{D}_k}, y_k), 1 \leq k \leq K \text{ such that} \\ & \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\ & \sum_{j \in \mathcal{J}} r_j > \sum_{j \in \mathcal{J}} H(V_j | V_{\mathcal{A}_j}) - H(V_{\mathcal{J}} | S) \\ & \text{for all } \mathcal{J} \subseteq [1 : N] \text{ such that } \mathcal{J} \neq \emptyset \text{ and if } j \in \mathcal{J}, \text{ then } \mathcal{A}_j \subseteq \mathcal{J}, \\ & \sum_{j \in \mathcal{J}^c} r_j < \sum_{j \in \mathcal{J}^c} H(V_j | V_{\mathcal{A}_j}) - H(V_{\mathcal{J}^c} | Y_k V_{\mathcal{J}}) \\ & \left. \text{for all } 1 \leq k \leq K \text{ and for all } \mathcal{J} \subseteq \mathcal{D}_k \text{ such that } \mathcal{J}^c \triangleq \mathcal{D}_k \setminus \mathcal{J} \neq \emptyset \text{ and if } j \in \mathcal{J}, \text{ then } \mathcal{A}_j \subseteq \mathcal{J} \right\}. \quad (11) \end{aligned}$$

²We say a set \mathcal{G} is larger than another \mathcal{H} if $|\mathcal{G}| > |\mathcal{H}|$, or $|\mathcal{G}| = |\mathcal{H}|$ and there exists some $1 \leq i \leq |\mathcal{G}|$ such that $\mathcal{G}[i] > \mathcal{H}[i]$ and $\mathcal{G}[l] = \mathcal{H}[l]$ for all $1 \leq l \leq i - 1$, where $\mathcal{G}[i]$ (or $\mathcal{H}[i]$) denotes the i th largest element in \mathcal{G} (or \mathcal{H}).

Besides, define two distortion region (outer bounds, achieved by introducing auxiliary random variables at sender)

$$\begin{aligned} \mathcal{R}_1^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmf } p_{\hat{S}_{[1:K]}|S} \text{ such that} \right. \\ & \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\ & \text{and for any pmf } p_{U_{[1:L]}|S}, \text{ one can find } p_{\tilde{U}_{[1:L]}, X^n} \text{ satisfying} \\ & \left. I(\hat{S}_{\mathcal{A}}; U_{\mathcal{B}} | U_{\mathcal{C}}) \leq \frac{1}{n} I(Y_{\mathcal{A}}^n; \tilde{U}_{\mathcal{B}} | \tilde{U}_{\mathcal{C}}) \text{ for any } \mathcal{A} \subseteq [1:K], \mathcal{B}, \mathcal{C} \subseteq [1:L] \right\}, \end{aligned} \quad (12)$$

$$\begin{aligned} \mathcal{R}_1'^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmf } p_{\hat{S}_{[1:K]}|S} \text{ such that} \right. \\ & \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\ & \text{and for any pmf } p_{U_{[1:L]}|S}, \text{ one can find } p_{X, \tilde{U}_{[1:L]}, W_{[1:K]}, W'_{[1:K]}} \text{ satisfying} \\ & \sum_{i=1}^m I(\hat{S}_{\mathcal{A}_i}; U_{\mathcal{B}_i} | U_{\cup_{j=0}^{i-1} \mathcal{B}_j}) \leq \sum_{i=1}^m I(Y_{\mathcal{A}_i}; \tilde{U}_{\mathcal{B}_i} \tilde{W}_{\mathcal{A}_{i+1}} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j} \tilde{W}_{\mathcal{A}_i} \tilde{W}_{\mathcal{A}_{i-1}}), \\ & \text{for any } m \geq 1, \mathcal{A}_i \subseteq [1:K], \mathcal{B}_i \subseteq [1:L], 0 \leq i \leq m, \mathcal{A}_0, \mathcal{A}_{m+1} \triangleq \emptyset, \\ & \text{and } \tilde{W}_{\mathcal{A}_i} \triangleq W_{\mathcal{A}_i}, \text{ if } i \text{ is odd; } W'_{\mathcal{A}_i}, \text{ otherwise, or } \tilde{W}_{\mathcal{A}_i} \triangleq W'_{\mathcal{A}_i}, \text{ if } i \text{ is odd; } W_{\mathcal{A}_i}, \text{ otherwise} \Big\}, \end{aligned} \quad (13)$$

and another distortion region (outer bound, achieved by introducing auxiliary random variables at receivers)

$$\begin{aligned} \mathcal{R}_2^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmf } p_X \text{ and some functions } \hat{s}_k^n(\tilde{y}_k), 1 \leq k \leq K \text{ such that} \right. \\ & \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\ & \text{and for any pmf } p_{U_{[1:L]}|Y_{[1:K]}}, \text{ one can find } p_{\tilde{Y}_{[1:K]}|S} p_{\tilde{U}_{[1:L]}|\tilde{Y}_{[1:K]}} \text{ satisfying} \\ & \left. I(S; \tilde{Y}_{\mathcal{B}} \tilde{U}_{\mathcal{B}'} | \tilde{Y}_{\mathcal{C}} \tilde{U}_{\mathcal{C}'}) \leq I(X; Y_{\mathcal{B}} U_{\mathcal{B}'} | Y_{\mathcal{C}} U_{\mathcal{C}'}) \text{ for any } \mathcal{B}, \mathcal{C} \subseteq [1:K], \mathcal{B}', \mathcal{C}' \subseteq [1:L] \right\}. \end{aligned} \quad (14)$$

Then we have the following theorem. The proof is given in Appendix C.

Theorem 1. For transmitting source S over general broadcast channel $p_{Y_{[1:K]}|X}$,

$$\mathcal{R}^{(i)} \subseteq \mathcal{R} \subseteq \mathcal{R}_1^{(o)} \cap \mathcal{R}_2^{(o)} \subseteq \mathcal{R}_1'^{(o)}. \quad (15)$$

Remark 1. The inner bound of Theorem 1 can be easily extended to Gaussian or any other well-behaved continuous-alphabet source-channel pair by standard discretization method [16, Thm. 3.3], and moreover for this case the outer bounds still hold. Theorem 1 can be also extended to the case of source-channel bandwidth mismatch, where m samples of a DMS are transmitted through n uses of a DM-BC. This can be accomplished by replacing the source and channel symbols in Theorem 1 by supersymbols of lengths m and n , respectively. Besides, Theorem 1 could be also extended to the problem of broadcasting correlated sources (by modifying the distortion measure) or source broadcast with channel input cost (by adding channel input constraint).

The inner bound $\mathcal{R}^{(i)}$ in Theorem 1 is achieved by a unified hybrid coding scheme depicted in Fig. 2. In this scheme, the codebook has a layered (or superposition) structure, and consists of randomly and independently

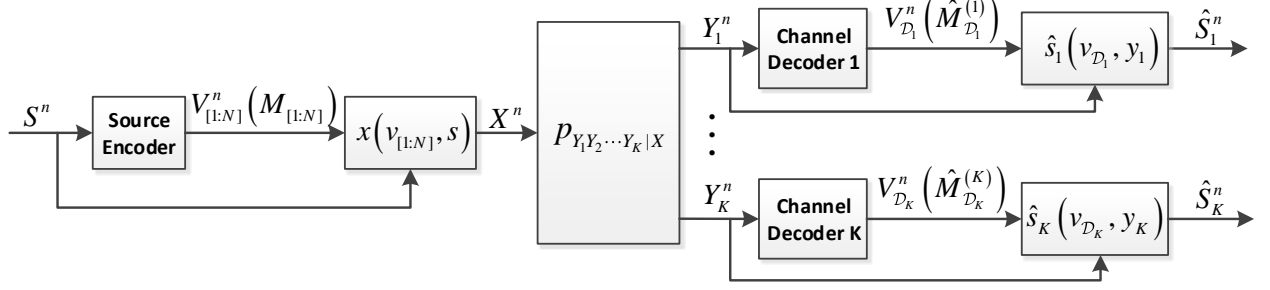


Fig. 2. A unified hybrid coding used to prove the inner bound of Theorem 1.

generated codewords $V_{[1:N]}^n(m_{[1:N]})$, $m_{[1:N]} \in \prod_{i=1}^N [1 : 2^{nr_i}]$, where $r_{[1:N]}$ satisfies (11). At encoder side, upon source sequence S^n , the encoder produces digital messages $M_{[1:N]}$ with M_i meant for all the receivers k satisfying $i \in \mathcal{D}_k$. Then, the codeword $V_{[1:N]}^n(M_{[1:N]})$ and the source sequence S^n are used to generate channel input X^n by symbol-by-symbol mapping $x(v_{[1:N]}, s)$. At decoder sides, upon received signal Y_k^n , decoder k could reconstruct $M_{\mathcal{D}_k}$ (and also $V_{\mathcal{D}_k}^n(M_{\mathcal{D}_k})$) losslessly, and then \hat{S}_k^n is produced by symbol-by-symbol mapping $\hat{s}_k(v_{\mathcal{D}_k}, y_k)$. Such a scheme could achieve any $D_{[1:K]}$ in the inner bound $\mathcal{R}^{(i)}$.

To reveal essence of such hybrid coding, the digital transmission part of this hybrid coding can be roughly understood as cascade of a K -user Gray-Wyner source-coding and a K -user Marton's broadcast channel-coding, which share a common codebook. According to [16, Thm. 13.3], the encoding operation of Gray-Wyner source-coding with rates $r_{[1:N]}$ is successful if $\sum_{j \in \mathcal{J}} r_j > \sum_{j \in \mathcal{J}} H(V_j | V_{\mathcal{A}_j}) - H(V_{\mathcal{J}} | S)$ for all $\mathcal{J} \subseteq [1 : N]$ such that $\mathcal{J} \neq \emptyset$ and if $j \in \mathcal{J}$, then $\mathcal{A}_j \subseteq \mathcal{J}$, and according to [16, Thm 5.2] the decoding operation of Marton's broadcast channel-coding with rates $r_{[1:N]}$ is successful if $\sum_{j \in \mathcal{J}^c} r_j < \sum_{j \in \mathcal{J}^c} H(V_j | V_{\mathcal{A}_j}) - H(V_{\mathcal{J}^c} | Y_k V_{\mathcal{J}})$ for all $1 \leq k \leq K$ and for all $\mathcal{J} \subseteq \mathcal{D}_k$ such that $\mathcal{J}^c \neq \emptyset$ and if $j \in \mathcal{J}$, then $\mathcal{A}_j \subseteq \mathcal{J}$. Note that in Marton's broadcast channel-coding, $r_{[1:N]}$ does not correspond to the regular broadcast-rates (i.e., the rates of subcodebooks in multicoding), but is the rates of the whole codebook. Since the proposed hybrid coding satisfies the two sufficient conditions above, $V_{\mathcal{D}_k}^n(M_{\mathcal{D}_k})$ could be losslessly transmitted to receiver k . Note that such informal understanding is inaccurate owing to the use of symbol-by-symbol mapping, but it provides the rationale for our scheme. Besides, the design of such unified hybrid coding is inspired by the hybrid coding scheme for sending correlated sources over multi-access channel in [15].

The outer bounds $\mathcal{R}_1^{(o)}$, $\mathcal{R}_1'^{(o)}$ and $\mathcal{R}_2^{(o)}$ of Theorem 1 are derived by introducing auxiliary random variables $U_{[1:L]}^n$ at sender side or at receiver sides. The proof method of introducing auxiliary random variables (or remote sources) at sender side could be found in [9], [11, Thm. 2] and [12, Lem. 1]. However, to the best of our knowledge, it is the first time to prove outer bounds by introducing auxiliary random variables (or remote channels) at receiver sides. In [9] it is used to derive the outer bound for Gaussian source broadcast, and in [11, Thm. 2] and [12, Lem. 1] it is used to derive the outer bounds for sending source over 2-user general broadcast channel. This proof

method generalizes the one used to derive trivial outer bound, but it does not always result in a tighter outer bound than the trivial one [10]. A deeper understanding of these proof methods has been given by Khezeli *et al.* in [12]. $p_{\hat{S}_{[1:K]}|S}$ can be considered as a virtual broadcast channel realized over physical broadcast channel $p_{Y_{[1:K]}|X}$, and hence certain measurements based on $p_{\hat{S}_{[1:K]}|S}$ are less than or equal to those based on $p_{Y_{[1:K]}|X}$. This leads to the necessary conditions on the communication. Besides, the necessary conditions can be also understood from the perspective of virtual source. X and $Y_{[1:K]}$ can be considered as a virtual source and K virtual reconstructions. Then the physical source S and the physical reconstructions $\hat{S}_{[1:K]}$ are correlated through the virtual source and virtual reconstructions. Hence the physical source should be more “tractable” than the virtual one, and certain measurements based on physical source and reconstructions should be less than or equal to those based on the virtual source and reconstructions. The analysis above gives the reasons why $\mathcal{R}_1^{(o)}$ and $\mathcal{R}_1'^{(o)}$ are expressed in form of comparison of the “capacity regions” of virtual broadcast channel and physical broadcast channel, while $\mathcal{R}_2^{(o)}$ is expressed in form of comparison of the “source-coding rate regions” of virtual source and physical source.

For 2-user broadcast case, the inner bound of Theorem 1 reduces to

$$\begin{aligned} \mathcal{R}^{(i)} = & \left\{ (D_1, D_2) : \text{There exist some pmf } p_{V_0, V_1, V_2|S}, \right. \\ & \text{and functions } x(v_0, v_1, v_2, s), \hat{s}_k(v_0, v_k, y_k), k = 1, 2, \text{ such that} \\ & \mathbb{E}d_k(S, \hat{S}_k) \leq D_k, \\ & I(V_0 V_k; S) < I(V_0 V_k; Y_k), k = 1, 2, \\ & I(V_0 V_1 V_2; S) + I(V_1; V_2|V_0) < \min\{I(V_0; Y_1), I(V_0; Y_2)\} + I(V_1; Y_1|V_0) + I(V_2; Y_2|V_0), \\ & \left. I(V_0 V_1; S) + I(V_0 V_2; S) + I(V_1; V_2|V_0 S) < I(V_0 V_1; Y_1) + I(V_0 V_2; Y_2) \right\}. \end{aligned} \quad (16)$$

This inner bound was first given in by Yassaee *et. al* [21]. On the other hand, for 2-user broadcast case, letting $L = 1$ for $\mathcal{R}_1^{(o)}$ and $\mathcal{R}_2^{(o)}$, and $L = 3$ for $\mathcal{R}_1'^{(o)}$, the outer bounds of Theorem 1 reduces to

$$\begin{aligned} \mathcal{R}_1^{(o)} = & \left\{ (D_1, D_2) : \text{There exists some pmf } p_{\hat{S}_1, \hat{S}_2|S} \text{ such that} \right. \\ & \mathbb{E}d_k(S, \hat{S}_k) \leq D_k, k = 1, 2, \\ & \text{and for any pmf } p_{U|S}, \text{ one can find } p_{\tilde{U}, X^n} \text{ satisfying} \\ & I(\hat{S}_1; U) \leq \frac{1}{n} I(Y_1^n; \tilde{U}), \\ & I(\hat{S}_2; U) \leq \frac{1}{n} I(Y_2^n; \tilde{U}), \\ & I(\hat{S}_1 \hat{S}_2; U) \leq \frac{1}{n} I(Y_1^n Y_2^n; \tilde{U}), \\ & I(\hat{S}_1; S|U) \leq \frac{1}{n} I(Y_1^n; X^n | \tilde{U}), \\ & I(\hat{S}_2; S|U) \leq \frac{1}{n} I(Y_2^n; X^n | \tilde{U}), \\ & \left. I(\hat{S}_1 \hat{S}_2; S|U) \leq \frac{1}{n} I(Y_1^n Y_2^n; X^n | \tilde{U}) \right\}, \end{aligned}$$

$\mathcal{R}_1^{(o)} = \left\{ (D_1, D_2) : \text{There exists some pmf } p_{\hat{S}_1, \hat{S}_2|S} \text{ such that} \right.$

$$\mathbb{E}d_k(S, \hat{S}_k) \leq D_k, k = 1, 2,$$

and for any pmf $p_{U_1, U_2, U_3|S}$, one can find $p_{X, \tilde{U}_1, \tilde{U}_2, \tilde{U}_3, W_1, W_2, W'_1, W'_2}$ satisfying

$$\begin{aligned} I(\hat{S}_{\mathcal{A}_1}; U_{\mathcal{B}_1} | U_{\mathcal{B}_0}) &\leq I(Y_{\mathcal{A}_1}; \tilde{U}_{\mathcal{B}_1} | \tilde{U}_{\mathcal{B}_0} \tilde{W}_{\mathcal{A}_1}), \\ I(\hat{S}_{\mathcal{A}_1}; U_{\mathcal{B}_1} | U_{\mathcal{B}_0}) + I(\hat{S}_{\mathcal{A}_2}; U_{\mathcal{B}_2} | U_{\mathcal{B}_0} U_{\mathcal{B}_1}) &\leq I(Y_{\mathcal{A}_1}; \tilde{U}_{\mathcal{B}_1} \tilde{W}_{\mathcal{A}_2} | \tilde{U}_{\mathcal{B}_0} \tilde{W}_{\mathcal{A}_1}) + I(Y_{\mathcal{A}_2}; \tilde{U}_{\mathcal{B}_2} | \tilde{U}_{\mathcal{B}_0} \tilde{U}_{\mathcal{B}_1} \tilde{W}_{\mathcal{A}_1} \tilde{W}_{\mathcal{A}_2}), \\ I(\hat{S}_{\mathcal{A}_1}; U_{\mathcal{B}_1} | U_{\mathcal{B}_0}) + I(\hat{S}_{\mathcal{A}_2}; U_{\mathcal{B}_2} | U_{\mathcal{B}_0} U_{\mathcal{B}_1}) + I(\hat{S}_{\mathcal{A}_3}; U_{\mathcal{B}_3} | U_{\mathcal{B}_0} U_{\mathcal{B}_1} U_{\mathcal{B}_2}) &\leq I(Y_{\mathcal{A}_1}; \tilde{U}_{\mathcal{B}_1} \tilde{W}_{\mathcal{A}_2} | \tilde{U}_{\mathcal{B}_0} \tilde{W}_{\mathcal{A}_1}) \\ &\quad + I(Y_{\mathcal{A}_2}; \tilde{U}_{\mathcal{B}_2} \tilde{W}_{\mathcal{A}_3} | \tilde{U}_{\mathcal{B}_0} \tilde{U}_{\mathcal{B}_1} \tilde{W}_{\mathcal{A}_1} \tilde{W}_{\mathcal{A}_2}) + I(Y_{\mathcal{A}_3}; \tilde{U}_{\mathcal{B}_3} | \tilde{U}_{\mathcal{B}_0} \tilde{U}_{\mathcal{B}_1} \tilde{U}_{\mathcal{B}_2} \tilde{W}_{\mathcal{A}_2} \tilde{W}_{\mathcal{A}_3}), \end{aligned}$$

for any $\mathcal{A}_i \subseteq [1 : 2], \mathcal{B}_i \subseteq [1 : 3], 0 \leq i \leq 3$,

$$\text{and } \tilde{W}_{\mathcal{A}_i} \triangleq W_{\mathcal{A}_i}, \text{ if } i \text{ is odd; } W'_{\mathcal{A}_i}, \text{ otherwise, or } \tilde{W}_{\mathcal{A}_i} \triangleq W'_{\mathcal{A}_i}, \text{ if } i \text{ is odd; } W_{\mathcal{A}_i}, \text{ otherwise} \}, \quad (17)$$

and

$\mathcal{R}_2^{(o)} = \left\{ (D_1, D_2) : \text{There exists some pmf } p_X \text{ and some functions } \hat{s}_k(\tilde{y}_k), k = 1, 2 \text{ such that} \right.$

$$\mathbb{E}d_k(S, \hat{S}_k) \leq D_k, k = 1, 2,$$

and for any pmf $p_{U|Y_1 Y_2}$, one can find $p_{\tilde{Y}_1 \tilde{Y}_2|S} p_{\tilde{U}|\tilde{Y}_1 \tilde{Y}_2}$ satisfying

$$\begin{aligned} I(S; \tilde{U}) &\leq I(X; U), \\ I(S; \tilde{Y}_1 | \tilde{U}) &\leq I(X; Y_1 | U), \\ I(S; \tilde{Y}_2 | \tilde{U}) &\leq I(X; Y_2 | U), \\ I(S; \tilde{Y}_1 | \tilde{Y}_2 \tilde{U}) &\leq I(X; Y_1 | Y_2 U), \\ I(S; \tilde{Y}_2 | \tilde{Y}_1 \tilde{U}) &\leq I(X; Y_2 | Y_1 U), \\ I(S; \tilde{Y}_1 \tilde{Y}_2 | \tilde{U}) &\leq I(X; Y_1 Y_2 | U). \end{aligned}$$

The outer bound $\mathcal{R}_1^{(o)}$ is as tight as if not tighter than that given by Khezeli *et al.* in [12, Thm. 1]. This is because the outer bound in [12, Lem 1] is just $\mathcal{R}_1^{(o)}$ with \mathcal{A}_i restricted to be an element (not a subset) of $[1 : 2]$. In addition, note that introducing nondegenerate auxiliary random variable(s) in the outer bounds $\mathcal{R}_1^{(o)}$, $\mathcal{R}_1'^{(o)}$ and $\mathcal{R}_2^{(o)}$ are not trivial in general. The necessity of introducing nondegenerate variable(s) for $\mathcal{R}_1^{(o)}$ and $\mathcal{R}_1'^{(o)}$ can be concluded from some special cases, e.g., source broadcast over degraded channel, quadratic Gaussian source broadcast or Hamming binary source broadcast (see the details in the subsequent three subsections). To show the necessity of introducing nondegenerate variable for $\mathcal{R}_2^{(o)}$, we only consider the first three inequalities on mutual information in $\mathcal{R}_2^{(o)}$. Next we show that the necessary conditions

$$I(S; \tilde{U}) \leq I(X; U), \quad (18)$$

$$I(S; \tilde{Y}_1 | \tilde{U}) \leq I(X; Y_1 | U), \quad (19)$$

$$I(S; \tilde{Y}_2 | \tilde{U}) \leq I(X; Y_2 | U), \quad (20)$$

with nondegenerate U results in a tighter bound than that with degenerate U (for the latter case, the necessary conditions reduce to the trivial bound).

Suppose the broadcast channel $P_{Y_1 Y_2 | X}$ satisfies $Y_1 = (Y_0, Y'_1), Y_2 = (Y_0, Y'_2)$ for some Y_0, Y'_1, Y'_2 . Suppose lossless transmission case (Hamming distortion measure and $D_1 = D_2 = 0$): $S = (S_1, S_2), \hat{S}_1 = S_1, \hat{S}_2 = S_2$, and $H(S_1) = C_1$ and $H(S_2) = C_2$. Hence the trivial outer bound implies (S_1, S_2) can be transmitted losslessly. Now we show that the outer bound $\mathcal{R}_2^{(o)}$ implies (S_1, S_2) can not be transmitted losslessly for some cases. Set $U = Y_0$ in $\mathcal{R}_2^{(o)}$. Then it is easy to obtain the following inequalities from (18)-(20).

$$I(S_1 S_2; V) \leq I(X; Y_0), \quad (21)$$

$$H(S_1 | V) \leq I(X; Y_1 | Y_0), \quad (22)$$

$$H(S_2 | V) \leq I(X; Y_2 | Y_0), \quad (23)$$

for some $p_{V|S_1 S_2}$. Therefore, we further have

$$H(S_1) \leq H(S_1) + I(S_2; V | S_1) = I(S_1 S_2; V) + H(S_1 | V) \leq I(X; Y_0) + I(X; Y_1 | Y_0) = I(X; Y_1) \leq C_1. \quad (24)$$

On the other hand, since $H(S_1) = C_1$ and $H(S_2) = C_2$, the equalities hold in the inequalities above, which implies $I(X; Y_1) = C_1$ (i.e., P_X is the capacity-achieving distribution), $I(S_1 S_2; V) = I(X; Y_0)$, and $I(S_2; V | S_1) = 0$, i.e., $S_2 \rightarrow S_1 \rightarrow V$. Similarly, we have $S_1 \rightarrow S_2 \rightarrow V$. In addition, the Gács-Körner common information [22], [23] can be expressed as

$$C_{GK}(S_1; S_2) = \sup_{P_{V|S_1 S_2}: S_2 \rightarrow S_1 \rightarrow V, S_1 \rightarrow S_2 \rightarrow V} I(S_1 S_2; V). \quad (25)$$

Hence there exists $p_{V|S_1 S_2}$ such that $S_2 \rightarrow S_1 \rightarrow V, S_1 \rightarrow S_2 \rightarrow V$ and $I(S_1 S_2; V) = I(X; Y_0)$, only if $C_{GK}(S_1; S_2) \geq I(X; Y_0) > 0$ (suppose the channel $P_{Y_0|X}$ satisfy $I(X; Y_0) > 0$ for the capacity-achieving distribution P_X). However GK common information does not always exist for any source pair (S_1, S_2) , e.g., $C_{GK}(S_1; S_2) = 0$ for doubly symmetric binary source. This implies the outer bound is tighter than the trivial one, which in turn implies the necessity of introducing nondegenerate variable for $\mathcal{R}_2^{(o)}$.

When consider lossless transmission of independent source, Theorem 1 can be used to achieve bounds on capacity region of general broadcast channel with common messages. In this case, $S = (M_j : 1 \leq j \leq N)$ and all M_j 's are independent with each other. For each $1 \leq j \leq N$, let R_j denote the rate of the common message M_j that is to be transmitted losslessly from sender to all the receivers in $\mathcal{G}_j \subseteq [1 : K]$. The correspondence between M_j and \mathcal{G}_j is kept same to that between V_j and \mathcal{G}_j ; see the beginning of this section. Then such achievable rates $R_{[1:N]}$ constitute the capacity region \mathcal{C} .

Now, define rate region

$$\begin{aligned}
\mathcal{C}^{(i)} = & \left\{ R_{[1:N]} : \text{There exist some pmf } p_{V_{[1:N]}|S}, \text{ vector } r_{[1:N]} \right. \\
& \text{and function } x(v_{[1:N]}) \text{ such that} \\
& \sum_{j \in \mathcal{J}} r_j > \sum_{j \in \mathcal{J}} R_j + \sum_{j \in \mathcal{J}} H(V_j|V_{\mathcal{A}_j}) - H(V_{\mathcal{J}}|S) \\
& \text{for all } \mathcal{J} \subseteq [1:N] \text{ such that } \mathcal{J} \neq \emptyset \text{ and if } j \in \mathcal{J}, \text{ then } \mathcal{A}_j \subseteq \mathcal{J}, \\
& \sum_{j \in \mathcal{J}^c} r_j < \sum_{j \in \mathcal{J}^c} H(V_j|V_{\mathcal{A}_j}) - H(V_{\mathcal{J}^c}|Y_k V_{\mathcal{J}}) \\
& \left. \text{for all } 1 \leq k \leq K \text{ and for all } \mathcal{J} \subseteq \mathcal{D}_k \text{ such that } \mathcal{J}^c \triangleq \mathcal{D}_k \setminus \mathcal{J} \neq \emptyset \text{ and if } j \in \mathcal{J}, \text{ then } \mathcal{A}_j \subseteq \mathcal{J} \right\}, \quad (26)
\end{aligned}$$

and another rate region

$$\begin{aligned}
\mathcal{C}^{(o)} = & \left\{ R_{[1:N]} : \text{There exists some pmf } \prod_{i=1}^N p_{\tilde{U}_i} p_{W_{[1:K]}, W'_{[1:K]} | \tilde{U}_{[1:N]}} \text{ and function } x(\tilde{u}_{[1:N]}) \text{ such that} \right. \\
& \sum_{i=1}^m \sum_{j \in \left(\bigcup_{k \in \mathcal{A}_i} \mathcal{D}_k \right) \cap \mathcal{B}_i} R_j \leq \sum_{i=1}^m I(Y_{\mathcal{A}_i}; \tilde{U}_{\mathcal{B}_i} \tilde{W}_{\mathcal{A}_{i+1}} | \tilde{U}_{\bigcup_{j=0}^{i-1} \mathcal{B}_j} \tilde{W}_{\mathcal{A}_i} \tilde{W}_{\mathcal{A}_{i-1}}), \\
& \text{for any } m \geq 1, \mathcal{A}_i \subseteq [1:K], \mathcal{B}_i \subseteq [1:N], 0 \leq i \leq m, \mathcal{A}_0, \mathcal{A}_{m+1} \triangleq \emptyset, \\
& \left. \text{and } \tilde{W}_{\mathcal{A}_i} \triangleq W_{\mathcal{A}_i}, \text{ if } i \text{ is odd; } W'_{\mathcal{A}_i}, \text{ otherwise, or } \tilde{W}_{\mathcal{A}_i} \triangleq W'_{\mathcal{A}_i}, \text{ if } i \text{ is odd; } W_{\mathcal{A}_i}, \text{ otherwise} \right\}. \quad (27)
\end{aligned}$$

Then as a consequence of Theorem 1, we can establish the following bounds on the capacity region of general broadcast channel. The proof is omitted.

Theorem 2. For general broadcast channel $p_{Y_{[1:K]}|X}$, the capacity region \mathcal{C} with common messages satisfies

$$\mathcal{C}^{(i)} \subseteq \mathcal{C} \subseteq \mathcal{C}^{(o)}. \quad (28)$$

Remark 2. The inner bound and the outer bound of Theorem 2 respectively follow from $\mathcal{R}^{(i)}$ and $\mathcal{R}_1'^{(o)}$ of Theorem 1, and respectively generalize Marton's inner bound and Nair-El Gamal outer bound to the case of K -user broadcast channel.

B. Source Broadcast over Degraded Channel

If the channel is degraded, define

$$\begin{aligned}
\mathcal{R}_{DBC}^{(i)} = & \left\{ D_{[1:K]} : \text{There exist some pmf } p_{V_K|S} p_{V_{K-1}|V_K} \cdots p_{V_1|V_2}, \right. \\
& \text{and functions } x(v_K, s), \hat{s}_k(v_k, y_k), 1 \leq k \leq K \text{ such that} \\
& \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, \\
& \left. I(S; V_k) \leq \sum_{j=1}^k I(Y_j; V_j | V_{j-1}), 1 \leq k \leq K, \text{ where } V_0 \triangleq \emptyset \right\}, \quad (29)
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{R}_{DBC}^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmfs } p_{\hat{S}_{[1:K]}|S}, p_X \text{ such that} \right. \\
& \mathbb{E}d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\
& \left(I(\hat{S}_{[1:k]}; U_k | U_{k-1}) : k \in [1:K] \right) \in \mathcal{B}_{DBC}(p_X p_{Y_{[1:K]}|X}) \\
& \text{for any pmf } p_{U_{K-1}|S} p_{U_{K-2}|U_{K-1}} \cdots p_{U_1|U_2}, U_0 \triangleq \emptyset, U_K \triangleq S, \\
& \left. \text{and } \left(I(S; \hat{S}_{[1:k]}) : k \in [1:K] \right) \in \mathcal{B}_{SRC}(p_X p_{Y_{[1:K]}|X}) \right\}, \tag{30}
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{B}_{DBC}(p_X p_{Y_{[1:K]}|X}) = & \left\{ R_{[1:K]} : \text{There exists some pmf } p_{V_{K-1}|X} p_{V_{K-2}|V_{K-1}} \cdots p_{V_1|V_2} \text{ such that} \right. \\
& R_k \geq 0, \sum_{j=1}^k R_j \leq \sum_{j=1}^k I(Y_j; V_j | V_{j-1}), 1 \leq k \leq K, \text{ where } V_0 \triangleq \emptyset, V_K \triangleq X \left. \right\} \tag{31}
\end{aligned}$$

denotes the capacity of degraded broadcast channel $p_{Y_{[1:K]}|X}$ with input X following p_X , and

$$\mathcal{B}_{SRC}(p_X p_{Y_{[1:K]}|X}) = \left\{ R_{[1:K]} : R_k \geq 0, \sum_{j=1}^k R_j \geq I(X; Y_{[1:k]}), 1 \leq k \leq K \right\} \tag{32}$$

denotes the successive refinement coding rate region of source X with reconstructions $Y_{[1:K]}$ following $p_{Y_{[1:K]}|X}$.

Then as a consequence of Theorem 1, the following theorem holds. The proof is given in Appendix D.

Theorem 3. For transmitting source S over degraded broadcast channel $p_{Y_{[1:K]}|X}$,

$$\mathcal{R}_{DBC}^{(i)} \subseteq \mathcal{R} \subseteq \mathcal{R}_{DBC}^{(o)}. \tag{33}$$

Remark 3. $\mathcal{R}_{DBC}^{(o)}$ can be also expressed as

$$\begin{aligned}
\mathcal{R}_{DBC}^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmf } p_{\hat{S}_{[1:K]}|S}, p_X \text{ such that} \right. \\
& \mathbb{E}d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\
& \mathcal{B}_{DBC}(p_S p_{\hat{S}'_{[1:K]}|S}) \subseteq \mathcal{B}_{DBC}(p_X p_{Y_{[1:K]}|X}), \\
& \left. \mathcal{B}_{SRC}(p_S p_{\hat{S}'_{[1:K]}|S}) \supseteq \mathcal{B}_{SRC}(p_X p_{Y_{[1:K]}|X}), \hat{S}'_k \triangleq \hat{S}_{[1:k]}, 1 \leq k \leq K \right\}. \tag{34}
\end{aligned}$$

From the last constraint of $\mathcal{R}_{DBC}^{(o)}$, one can obtain an interesting result: the trivial outer bound $D_{[1:K]}^*$ can be achieved for source broadcast over degraded channel only if the source is successively refinable.

Note that the last constraint of $\mathcal{R}_{DBC}^{(i)}$ can be understood as the intersection between the rate region of successive refinement coding for source S and reconstructions $V_{[1:K]}$ and the capacity of degraded broadcast channel $p_{Y_{[1:K]}|X}$ with input X and auxiliary random variables $V_{[1:K]}$, is not empty. The second constraint of $\mathcal{R}_{DBC}^{(o)}$ can be understood as the capacity of virtual degraded broadcast channel $p_{\hat{S}'_{[1:K]}|S}$ with input S is included in the capacity of physical degraded broadcast channel $p_{Y_{[1:K]}|X}$ with input X . The last constraint of $\mathcal{R}_{DBC}^{(o)}$ can be understood as the rate region of successive refinement coding for the physical source S and reconstructions $\hat{S}'_{[1:K]}$ includes the rate region

of successive refinement coding for the virtual source X and reconstructions $Y_{[1:K]}$. Similar to the general broadcast channel case, these necessary conditions is also consistent with the intuition. From the perspective of channel, the virtual broadcast channel is realized over the physical broadcast channel, hence the physical channel should be more “capable” than the virtual one. On the other hand, from the perspective of source, the physical source is correlated with the reconstructions through the virtual source and virtual reconstructions, hence the physical source should be more “tractable” than the virtual one.

C. Quadratic Gaussian Source Broadcast

Consider sending Gaussian source $S \sim \mathcal{N}(0, N_S)$ with quadratic distortion measure $d_k(s, \hat{s}) = d(s, \hat{s}) \triangleq (s - \hat{s})^2$ over power-constrained Gaussian broadcast channel $Y_k = X + W_k, 1 \leq k \leq K$ with $\mathbb{E}(X^2) \leq P$ and $W_k \sim \mathcal{N}(0, N_k), N_1 \geq N_2 \geq \dots \geq N_K$. Assume bandwidth mismatch factor is b . Then the inner bound $\mathcal{R}_{DBC}^{(i)}$ in Theorem 3 could recover the best known inner bound so far [7, Thm. 5] by setting suitable random variables and symbol-by-symbol mappings.

Corollary 1. [7, Thm. 5] For transmitting Gaussian source S with quadratic distortion measure over 2-user Gaussian broadcast channel with bandwidth mismatch factor b , $\mathcal{R}_{DBC}^{(i)} \subseteq \mathcal{R}$.

- For $b < 1$ (bandwidth compression)

$$\mathcal{R}_{DBC}^{(i)} = \{(D_1(\lambda, \gamma), D_2(\lambda, \gamma)) : 0 \leq \lambda \leq 1, 0 \leq \gamma \leq 1\}, \quad (35)$$

where

$$D_1(\lambda, \gamma) = \frac{bN_S}{\frac{\lambda P + N_1}{\lambda \gamma P + N_1}} + \frac{(1-b)N_S}{\left(\frac{P+N_1}{\lambda P + N_1}\right)^{\frac{b}{1-b}}}, \quad (36)$$

$$D_2(\lambda, \gamma) = \frac{bN_S}{\frac{\lambda P + N_1}{\lambda \gamma P + N_1}} + \frac{(1-b)N_S}{\left(\frac{P+N_1}{\lambda P + N_1} \frac{\lambda \gamma P + N_2}{N_2}\right)^{\frac{b}{1-b}}}. \quad (37)$$

- For $b > 1$ (bandwidth expansion)

$$\mathcal{R}_{DBC}^{(i)} = \{(D_1(\lambda, \gamma), D_2(\lambda, \gamma)) : 0 \leq \lambda \leq 1, 0 \leq \gamma \leq 1\}, \quad (38)$$

where

$$D_1(\lambda, \gamma) = \frac{N_S}{\left(\frac{\frac{b(1-\gamma)}{b-1}P + N_1}{\lambda \frac{b(1-\gamma)}{b-1}P + N_1}\right)^{b-1} \left(\frac{b\gamma P + N_1}{N_1}\right)}, \quad (39)$$

$$D_2(\lambda, \gamma) = \frac{N_S}{\left(\frac{\frac{b(1-\gamma)}{b-1}P + N_1}{\lambda \frac{b(1-\gamma)}{b-1}P + N_1}\right)^{b-1} \left(\frac{b\gamma P + N_2}{N_2}\right) \left(\frac{\lambda \frac{b(1-\gamma)}{b-1}P + N_2}{N_2}\right)^{b-1}}. \quad (40)$$

Proof: For $b = \frac{n}{m} < 1$ (bandwidth compression), the source $S^m = (S^n, S_{n+1}^m)$ is transmitted over channel $Y_k^n = X^n + W_k^n, k = 1, 2$. Define a set of random variables $(U_1^{m-n}, E_1^{m-n}, U_2^{m-n}, E_2^{m-n}, X_1^n, X_2^n, X_2'^n, V_1, V_2)$

such that

$$S_{n+1}^m = U_1^{m-n} + E_1^{m-n}, \quad (41)$$

$$E_1^{m-n} = U_2^{m-n} + E_2^{m-n}, \quad (42)$$

$$X_2^n = X_2'^n + \beta \alpha S^n, \quad (43)$$

$$V_1 = (U_1^{m-n}, X_1^n), \quad (44)$$

$$V_2 = (V_1, U_2^{m-n}, X_2^n), \quad (45)$$

where $U_1^{m-n}, U_2^{m-n}, E_2^{m-n}$ are mutually independent Gaussian variables, $X_2'^n$ and X_1^n are Gaussian variables independent of all the other variables, $\text{Var}(X_2') = \lambda \gamma P$, $\text{Var}(X_1) = (1 - \lambda) P$, $\text{Var}(E_1) = \frac{N_S}{\left(\frac{P+N_1}{\lambda P+N_1}\right)^{\frac{b}{1-b}}}$, $\text{Var}(E_2) = \frac{\text{Var}(E_1)}{\left(\frac{\lambda \gamma P+N_2}{N_2}\right)^{\frac{b}{1-b}}}$, and $\alpha = \sqrt{\frac{\lambda(1-\gamma)P}{N_S}}$, $\beta = \frac{\lambda \gamma P}{\lambda \gamma P+N_2}$. Define a set of functions

$$x^n(v_2, s^m) = x_1^n + \alpha s^n + x_2^n - \beta \alpha s^n = x_1^n + \alpha s^n + x_2'^n, \quad (46)$$

$$\hat{s}_1^m(v_1, y_1^n) = \left(\frac{\alpha N_S}{\alpha^2 N_S + N_1} (y_1^n - x_1^n), u_1^{m-n} \right), \quad (47)$$

$$\hat{s}_2^m(v_2, y_2^n) = \left(\frac{\alpha N_S}{\alpha^2 N_S + N_2} (y_2^n - x_1^n), u_1^{m-n} + u_2^{m-n} \right). \quad (48)$$

Substitute these variables and functions into the inner bound $\mathcal{R}_{DBC}^{(i)}$ in Theorem 3, then the $b < 1$ case in Corollary 1 is recovered.

For $b = \frac{n}{m} > 1$ (bandwidth expansion), the source S^n is transmitted over channel $(Y_k^m, Y_{k,m+1}^n) = (X^m, X_{m+1}^n) + (W_k^m, W_{k,m+1}^n)$, $k = 1, 2$. Define a set of random variables $(U_1^m, E_1^m, U_2^m, E_2^m, X_1^{n-m}, X_2^{n-m}, V_1, V_2)$ such that

$$S^m = U_1^m + E_1^m, \quad (49)$$

$$E_1^m = U_2^m + E_2^m, \quad (50)$$

$$V_1 = (U_1^m, X_1^{n-m}), \quad (51)$$

$$V_2 = (V_1, U_2^m, X_2^{n-m}), \quad (52)$$

where U_1^m, U_2^m, E_2^m are mutually independent Gaussian variables, X_1^{n-m} and X_2^{n-m} are two Gaussian variables independent of all the other random variables, $\text{Var}(X_1) = \frac{(1-\lambda)(1-\gamma)bP}{b-1}$, $\text{Var}(X_2) = \frac{\lambda(1-\gamma)bP}{b-1}$, $\text{Var}(E_1) = \frac{N_S}{\left(\frac{b(1-\gamma)P+N_1}{\lambda \frac{b(1-\gamma)P+N_1}{b-1}}\right)^{b-1}}$ and $\text{Var}(E_2) = \frac{\text{Var}(E_1)}{\left(\frac{b\gamma P+N_2}{N_2}\right)\left(\frac{\lambda \frac{b(1-\gamma)P+N_2}{b-1}}{N_2}\right)^{b-1} - \frac{b\gamma P}{N_2}}$. Define a set of functions

$$x^n(v_2, s^m) = (\alpha(s^m - u_1^m), x_1^{n-m} + x_2^{n-m}) = (\alpha e_1^m, x_1^{n-m} + x_2^{n-m}), \quad (53)$$

$$\hat{s}_1^m(v_1, y_1^n) = u_1^m + \frac{\alpha \text{Var}(E_1)}{\alpha^2 \text{Var}(E_1) + N_1} y_1^m, \quad (54)$$

$$\hat{s}_2^m(v_2, y_2^n) = u_1^m + \frac{N_2}{\alpha^2 \text{Var}(E_2) + N_2} u_2^m + \frac{\alpha \text{Var}(E_2)}{\alpha^2 \text{Var}(E_2) + N_2} y_2^m, \quad (55)$$

where $\alpha = \sqrt{\frac{\gamma b P}{\text{Var}(E_1)}}$. Substitute these variables and functions into the inner bound $\mathcal{R}_{DBC}^{(i)}$ in Theorem 3, then the $b > 1$ case in Corollary 1 is also recovered. \blacksquare

On the other hand, setting $U_{[1:K-1]}$ to be jointly Gaussian with S , the outer bound $\mathcal{R}_{DBC}^{(o)}$ in Theorem 3 could recover the best known outer bound [9, Thm. 2].

Theorem 4. [9, Thm. 2] For transmitting Gaussian source S with quadratic distortion measure over K -user Gaussian broadcast channel with bandwidth mismatch factor b ,

$$\mathcal{R} \subseteq \mathcal{R}_{DBC}^{(o)} \triangleq \left\{ D_{[1:K]} : \text{For any variables } +\infty = \tau_0 \geq \tau_1 \geq \dots \geq \tau_K = 0, \right. \\ \left. \frac{1}{b} \left(\frac{1}{2} \log \frac{(N_S + \tau_k)(D_k + \tau_{k-1})}{(D_k + \tau_k)(N_S + \tau_{k-1})} : k \in [1:K] \right) \in \mathcal{C}_{GBC} \right\}, \quad (56)$$

where \mathcal{C}_{GBC} denotes the capacity of Gaussian broadcast channel given by

$$\mathcal{C}_{GBC} = \left\{ R_{[1:K]} : R_k \geq 0, 1 \leq k \leq K, N_{K+1} = 0, \sum_{k=1}^K (N_k - N_{k+1}) \exp \left(2 \sum_{j=1}^k R_j \right) \leq P + N_1 \right\}. \quad (57)$$

To compare $\mathcal{R}_{DBC}^{(o)}$ with the trivial distortion bound, we consider a set of Gaussian test channels $p_{\hat{S}_k^*|S}$, $1 \leq k \leq K$ that achieves the optimal point-to-point distortion for each receiver. Assume $p_{\hat{S}_{[1:K]}^*|S}$ is the backward Gaussian broadcast channel consisting of subchannels $p_{\hat{S}_k^*|S}$, $1 \leq k \leq K$. Since any backward Gaussian channel can be transformed into a forward Gaussian channel with the same probability distribution, assume $p_{V_{[1:K]}|U}$ is the forward Gaussian broadcast channel of $p_{\hat{S}_{[1:K]}^*|S}$. Then for Gaussian source broadcast, the outer bound is in form of the comparison of capacity regions for two Gaussian broadcast channels $p_{V_{[1:K]}|U}$ and $p_{Y_{[1:K]}|X}$. Note that $p_{Y_{[1:K]}|X}$ and $p_{V_{[1:K]}|U}$ have different bandwidth (the bandwidth ratio is b) but the same point-to-point capacity for each receiver. It can be proved that the Gaussian broadcast capacity region shrinks as the bandwidth increases. Additionally, $\mathcal{C}_{DBC}(p_{V_{[1:K]}|U}) = \mathcal{C}_{DBC}(p_{Y_{[1:K]}|X})$ when bandwidth matched. Hence $\mathcal{C}_{DBC}(p_{V_{[1:K]}|U}) \subseteq \mathcal{C}_{DBC}(p_{Y_{[1:K]}|X})$ always holds for bandwidth compression case. This is the reason why the outer bound in [9] is nontrivial only for bandwidth expansion. The details can be found in [10].

The bounds in Corollary 1 and Theorem 4 are illustrated in Fig. 3.

D. Hamming Binary Source Broadcast

Consider sending binary source $S \sim \text{Bern}(\frac{1}{2})$ with Hamming distortion measure $d_k(s, \hat{s}) = d(s, \hat{s}) \triangleq 0$, if $s = \hat{s}$; 1, otherwise, over binary broadcast channel $Y_k = X \oplus W_k$, $1 \leq k \leq K$ with $W_k \sim \text{Bern}(p_k)$, $\frac{1}{2} \geq p_1 \geq p_2 \geq \dots \geq p_K \geq 0$. Assume bandwidth mismatch factor is b .

We first consider the inner bound. For bandwidth expansion ($b > 1$), as a special case of hybrid coding systematic source-channel coding (Uncoded Systematic Coding) has been first investigated in [6]. For any point-to-point lossless communication system, such systematic coding does not loss the optimality; however, for some lossy cases such as Hamming binary source communication, it is not optimal any more. To retain the optimality, we can first quantize the source S , and then transmit the quantized signal using Uncoded Systematic Coding. The performance of this code could be obtained directly from Theorem 3.

Specifically, let $U_2 = S \oplus E_2$, $U_1 = U_2 \oplus E_1$ with $E_2 \sim \text{Bern}(D_2)$, $E_1 \sim \text{Bern}(d_1)$. Let $V_2 = (U_2, X^{b-1})$, $V_1 = (U_1, X_1^{b-1})$, $X_1^{b-1} = X^{b-1} \oplus B^{b-1}$, where X_1^{b-1} and X^{b-1} are independent of U_2 and U_1 , and X^{b-1} and B^{b-1}

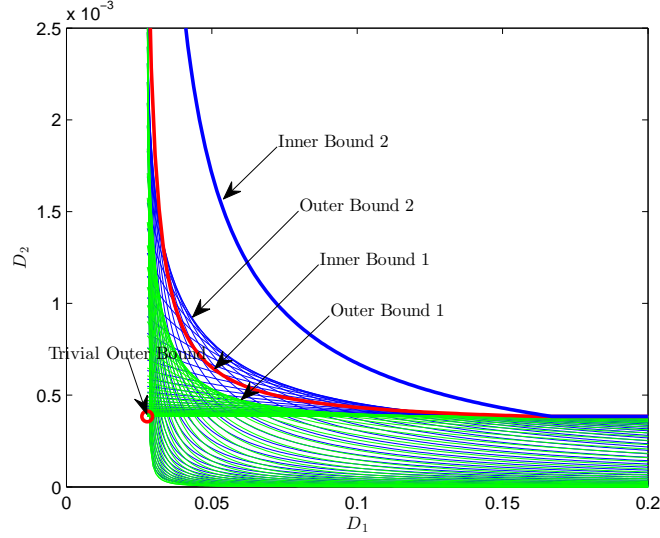


Fig. 3. Distortion bounds for sending Gaussian source over Gaussian broadcast channel with $b = 2$, $N_S = 1$, $P = 50$, $N_1 = 10$, $N_2 = 1$. Outer Bounds 1 and 2 and Inner Bounds 1 and 2 respectively correspond to the outer bound of Theorem 4, the outer bound of Theorem 8, the inner bound in Corollary 1, and the inner bound achieved by Wyner-Ziv separate coding (uncoded systematic code) [17, Lem. 3]. Trivial Outer Bound corresponds to the trivial outer bound (6). Besides, Outer Bound 2 and Inner Bound 2 could be considered as the outer bound and inner bound for Wyner-Ziv source broadcast problem with $b = 1$, $\beta_1 = \frac{N_S N_1}{P + N_1}$, $\beta_2 = \frac{N_S N_2}{P + N_2}$, and in this case Trivial Outer Bound corresponds to the Wyner-Ziv outer bound (65).

follow $b - 1$ dimensional $\text{Bern}(\frac{1}{2})$ and $\text{Bern}(\theta)$, respectively. Let $x^b(v_2, s) = (u_2, x^{b-1})$, $\hat{s}_2(v_2, y_2^b) = u_2$ and $\hat{s}_1(v_1, y_1^b) = u_1$, if $d_1 < p_1$; y_1 , otherwise. Substitute these variables and functions into the inner bound $\mathcal{R}_{DBC}^{(i)}$ in Theorem 3, then we get the following corollary.

Corollary 2 (Coded Systematic Coding). *For transmitting binary source S with Hamming distortion measure over 2-user binary broadcast channel with bandwidth mismatch factor b ,*

$$\begin{aligned}
 \mathcal{R} \supseteq \mathcal{R}_{CSC}^{(i)} &\triangleq \text{convexhull} \left\{ (D_1, D_2) : 0 \leq \theta, d_1 \leq \frac{1}{2}, \right. \\
 &\quad D_1 \geq \min \{d_1 \star D_2, p_1 \star D_2\}, \\
 &\quad r_1 = 1 - H_2(d_1 \star p_1) + (b - 1) [1 - H_2(\theta \star p_1)], \\
 &\quad r_2 = H_2(d_1 \star p_2) - H_2(p_2) + (b - 1) [H_2(\theta \star p_2) - H_2(p_2)], \\
 &\quad 1 - H_2(d_1 \star D_2) \leq r_1, \\
 &\quad \left. 1 - H_2(D_2) \leq r_1 + r_2 \right\}, \tag{58}
 \end{aligned}$$

where \star denotes the binary convolution, i.e.,

$$x \star y = (1 - x)y + x(1 - y), \tag{59}$$

and H_2 denotes the binary entropy function, i.e.,

$$H_2(p) = -p \log p - (1-p) \log(1-p). \quad (60)$$

Remark 4. Coded Systematic Coding without timesharing does not always lead to a convex distortion region, hence a timesharing mechanism is needed to improve performance and achieve $\mathcal{R}_{CSC}^{(i)}$. It is equivalent to adding a timesharing variable Q into V_2 and V_1 , before substitute them into the inner bound $\mathcal{R}_{DBC}^{(i)}$. Besides, note that unlike Uncoded Systematic Coding, the Coded Systematic Coding could always achieve the optimal distortion for at least one of the receivers. Moreover, unlike separate coding the Coded Systematic Coding could weaken the cliff effect, and result in slope-cliff effect.

In addition, the outer bound of Theorem 3 reduces to the following outer bound for Hamming binary source broadcast problem. This outer bound was first given in [11, Eqn (41)] for 2-user case. The proof is similar to that of [11, Eqn (41)], hence it is omitted here.

Theorem 5. For transmitting binary source S with Hamming distortion measure over K -user binary broadcast channel with bandwidth mismatch factor b ,

$$\mathcal{R} \subseteq \mathcal{R}_{DBC}^{(o)} \triangleq \left\{ D_{[1:K]} : \text{For any variables } \frac{1}{2} = \tau_0 \geq \tau_1 \geq \tau_2 \geq \dots \geq \tau_K = 0, \right. \\ \left. \frac{1}{b} (H_2(\tau_{k-1} \star D_k) - H_2(\tau_k \star D_k)) : k \in [1:K] \right\} \in \mathcal{C}_{BBC}, \quad (61)$$

where \mathcal{C}_{BBC} denotes the capacity of binary broadcast channel given by

$$\mathcal{C}_{BBC} = \left\{ R_{[1:K]} : \text{There exist some variables } \frac{1}{2} = \theta_0 \geq \theta_1 \geq \theta_2 \geq \dots \geq \theta_K = 0 \text{ such that} \right. \\ \left. 0 \leq R_k \leq H_2(\theta_{k-1} \star p_k) - H_2(\theta_k \star p_k), 1 \leq k \leq K \right\}. \quad (62)$$

The bounds in Corollary 2 and Theorem 5 are illustrated in Fig. 4.

IV. WYNER-ZIV SOURCE BROADCAST: SOURCE BROADCAST WITH SIDE INFORMATION

We now extend the problem by allowing decoders to access side information correlated with the source. As depicted in Fig. 5, receiver k observes memoryless side information Z_k^n , and it produces source reconstruction \hat{S}_k^n from received signal Y_k^n and side information Z_k^n .

Definition 7. An n -length Wyner-Ziv source-channel code is defined by the encoding function $x^n : \mathcal{S}^n \mapsto \mathcal{X}^n$ and a sequence of decoding functions $\hat{s}_k^n : \mathcal{Y}_k^n \times \mathcal{Z}_k^n \mapsto \hat{\mathcal{S}}_k^n, 1 \leq k \leq K$.

Definition 8. If there exists a sequence of Wyner-Ziv source-channel codes satisfying

$$\limsup_{n \rightarrow \infty} \mathbb{E} d_k \left(S^n, \hat{S}_k^n \right) \leq D_k, \quad (63)$$

then we say that the distortion tuple $D_{[1:K]}$ is achievable.

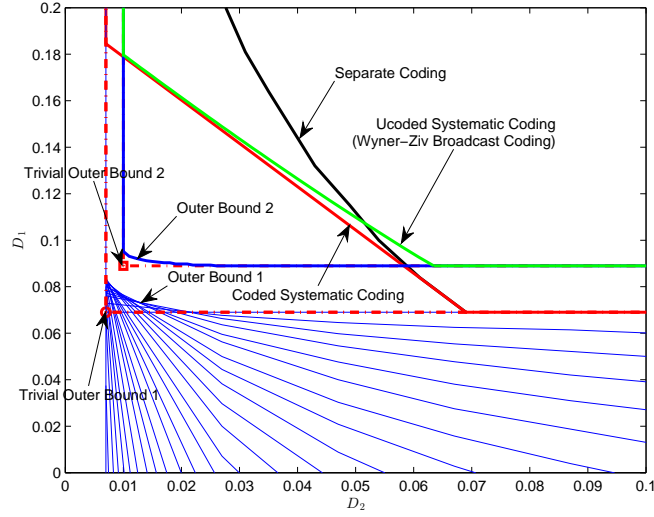


Fig. 4. Distortion bounds for sending binary source over binary broadcast channel with $b = 2, p_1 = 0.18, p_2 = 0.12$. Outer Bounds 1 and 2 respectively correspond to the outer bound of Theorem 5 and the outer bound of Theorem 9. Separate Coding, Uncoded Systematic Coding and Coded Systematic Coding respectively correspond to the separate scheme combining successive-refinement code [16, Example 13.3] with superposition code [16, Example 5.3], the inner bound in Corollary 3, and the inner bound in Corollary 2. Trivial Outer Bounds 1 and 2 correspond to the trivial outer bound (6) and the Wyner-Ziv outer bound (65), respectively. Besides, Trivial Outer Bound 2, Outer Bound 2 and Uncoded Systematic Coding could be considered as the outer bounds and inner bound for Wyner-Ziv source broadcast problem with $b = 1, \beta_1 = p_1, \beta_2 = p_2$.

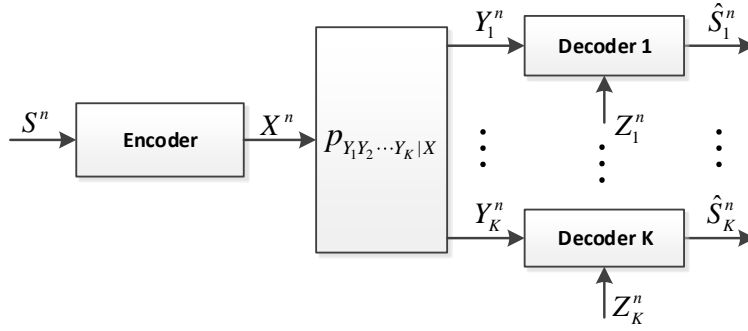


Fig. 5. Wyner-Ziv source broadcast system: broadcast communication system with side information at decoders.

Definition 9. The admissible distortion region for Wyner-Ziv broadcast problem is defined as

$$\mathcal{R}_{\text{SI}} \triangleq \{D_{[1:K]} : D_{[1:K]} \text{ is achievable}\}. \quad (64)$$

In addition, Shamai *et al.* [6, Thm. 2.1] showed that for transmitting source over point-to-point channel $p_{Y_k|X}$ with side information Z_k available at decoder, the minimum achievable distortion satisfies $R_{S|Z_k}(D_k) = C_k$, where $R_{S|Z_k}(\cdot)$ is the Wyner-Ziv rate-distortion function of the source S given that the decoder observes Z_k [16].

Therefore, the optimal distortion is $D_{\text{SI},k}^* = R_{S|Z_k}^{-1}(C_k)$. Obviously,

$$\mathcal{R}_{\text{SI}} \subseteq \mathcal{R}_{\text{SI}}^* \triangleq \{D_{[1:K]} : D_k \geq D_{\text{SI},k}^*, 1 \leq k \leq K\}, \quad (65)$$

where $\mathcal{R}_{\text{SI}}^*$ is named *Wyner-Ziv outer bound*.

Besides, we also consider the communication system with bandwidth mismatch, whereby m samples of a DMS are transmitted through n uses of a DM-BC with l samples of side information available at each decoder. For simplicity, we let $m = l$, and for this case, bandwidth mismatch factor is defined as $b = \frac{n}{m}$.

A. Wyner-Ziv Source Broadcast

If consider $Z_{[1:K]}$ to be transmitted from sender to the receivers over a virtual broadcast channel $p_{Z_{[1:K]}|S}$, and define $X' = (S, X)$ and $Y'_k = (Z_k, Y_k)$, $1 \leq k \leq K$, then the Wyner-Ziv source broadcast problem is equivalent to the problem of sending p_S over $p_{Y'_{[1:K]}|X'}$ with S restricted to be the input of subchannel $p_{Z_{[1:K]}|S}$. Hence Wyner-Ziv source broadcast problem could be considered as the problem of (uncoded) systematic source-channel coding. If set $x'(v_{[1:N]}, s) = (s, x(v_{[1:N]}, s))$, then from Theorem 1, we obtain the following inner bound for such systematic source-channel coding problem. It is therefore also an inner bound for the Wyner-Ziv source broadcast problem.

$$\begin{aligned} \mathcal{R}_{\text{SI}}^{(i)} = & \left\{ D_{[1:K]} : \text{There exist some pmf } p_{V_{[1:N]}|S}, \text{ vector } r_{[1:N]}, \right. \\ & \text{and functions } x(v_{[1:N]}, s), \hat{s}_k(v_{\mathcal{D}_k}, y_k, z_k), 1 \leq k \leq K \text{ such that} \\ & \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\ & \sum_{j \in \mathcal{J}} r_j > \sum_{j \in \mathcal{J}} H(V_j|V_{\mathcal{A}_j}) - H(V_{\mathcal{J}}|S) \\ & \text{for all } \mathcal{J} \subseteq [1:N] \text{ such that } \mathcal{J} \neq \emptyset \text{ and if } j \in \mathcal{J}, \text{ then } \mathcal{A}_j \subseteq \mathcal{J}, \\ & \sum_{j \in \mathcal{J}^c} r_j < \sum_{j \in \mathcal{J}^c} H(V_j|V_{\mathcal{A}_j}) - H(V_{\mathcal{J}^c}|Y_k Z_k V_{\mathcal{J}}) \\ & \left. \text{for all } 1 \leq k \leq K \text{ and for all } \mathcal{J} \subseteq \mathcal{D}_k \text{ such that } \mathcal{J}^c \triangleq \mathcal{D}_k \setminus \mathcal{J} \neq \emptyset \text{ and if } j \in \mathcal{J}, \text{ then } \mathcal{A}_j \subseteq \mathcal{J} \right\}. \quad (66) \end{aligned}$$

In addition, regard $(U_{[1:L]}, Z_{[1:K]})$ as auxiliary random variables following $p_{U_{[1:L]}|S} p_{Z_{[1:K]}|S}$, then following similar steps to the proof of the outer bounds $\mathcal{R}_1^{(o)}$ and $\mathcal{R}'_1^{(o)}$ of Theorem 1, we can achieve the following two outer bounds on \mathcal{R}_{SI} .

$$\begin{aligned} \mathcal{R}_{\text{SI},1}^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmf } p_{\hat{S}_{[1:K]}|S, Z_{[1:K]}} \text{ such that} \right. \\ & \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\ & \text{and for any pmf } p_{U_{[1:L]}|S}, \text{ one can find } p_{\tilde{U}_{[1:L]}, \tilde{Z}_{[1:K]}, X^n} \text{ satisfying} \\ & I(\hat{S}_{\mathcal{A}}; U_{\mathcal{B}} | U_{\mathcal{C}} Z_{\mathcal{A}}) \leq \frac{1}{n} I(Y_{\mathcal{A}}^n; \tilde{U}_{\mathcal{B}} | \tilde{U}_{\mathcal{C}} \tilde{Z}_{\mathcal{A}}) \text{ for any } \mathcal{A} \subseteq [1:K], \mathcal{B}, \mathcal{C} \subseteq [1:L] \left. \right\}, \quad (67) \end{aligned}$$

and

$$\begin{aligned}
\mathcal{R}_{\text{SI},1}'^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmf } p_{\hat{S}_{[1:K]}|S, Z_{[1:K]}} \text{ such that} \right. \\
& \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\
& \text{and for any pmf } p_{U_{[1:L]}|S}, \text{ one can find } p_{X, \tilde{U}_{[1:L]}, \tilde{Z}_{[1:K]}, W_{[1:K]}, W'_{[1:K]}} \text{ satisfying} \\
& \sum_{i=1}^m I(\hat{S}_{\mathcal{A}_i}; U_{\mathcal{B}_i} Z_{\mathcal{A}_{i+1}} | U_{\cup_{j=0}^{i-1} \mathcal{B}_j} Z_{\cup_{j=1}^i \mathcal{A}_j}) \leq \sum_{i=1}^m I(Y_{\mathcal{A}_i}; \tilde{U}_{\mathcal{B}_i} \tilde{Z}_{\mathcal{A}_{i+1}} \tilde{W}_{\mathcal{A}_{i+1}} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j} \tilde{Z}_{\cup_{j=1}^i \mathcal{A}_j} \tilde{W}_{\mathcal{A}_i} \tilde{W}_{\mathcal{A}_{i-1}}), \\
& \text{for any } m \geq 1, \mathcal{A}_i \subseteq [1:K], \mathcal{B}_i \subseteq [1:L], 0 \leq i \leq m, \mathcal{A}_0, \mathcal{A}_{m+1} \triangleq \emptyset, \\
& \left. \text{and } \tilde{W}_{\mathcal{A}_i} \triangleq W_{\mathcal{A}_i}, \text{ if } i \text{ is odd; } W'_{\mathcal{A}_i}, \text{ otherwise, or } \tilde{W}_{\mathcal{A}_i} \triangleq W'_{\mathcal{A}_i}, \text{ if } i \text{ is odd; } W_{\mathcal{A}_i}, \text{ otherwise} \right\}. \quad (68)
\end{aligned}$$

Following similar steps to the proof of the outer bound $\mathcal{R}_2^{(o)}$ of Theorem 1, we can also prove the following outer bound on \mathcal{R}_{SI} .

$$\begin{aligned}
\mathcal{R}_{\text{SI},2}^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmf } p_X \text{ and some functions } \hat{s}_k^n(\tilde{y}_k, z_k^n), 1 \leq k \leq K \text{ such that} \right. \\
& \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\
& \text{and for any pmf } p_{U_{[1:L]}|Y_{[1:K]}}, \text{ one can find } p_{\tilde{Y}_{[1:K]}|S, \tilde{U}_{[1:L]}|Y_{[1:K]}} \text{ satisfying} \\
& I(S^n; \tilde{Y}_{\mathcal{B}} \tilde{U}_{\mathcal{B}'} | \tilde{Y}_{\mathcal{C}} \tilde{U}_{\mathcal{C}'}) \leq I(X; Y_{\mathcal{B}} U_{\mathcal{B}'} | Y_{\mathcal{C}} U_{\mathcal{C}'}), \text{ for any } \mathcal{B}, \mathcal{C} \subseteq [1:K], \mathcal{B}', \mathcal{C}' \subseteq [1:L] \left. \right\}. \quad (69)
\end{aligned}$$

Therefore, the following theorem holds. The proof is omitted.

Theorem 6. For transmitting source S over broadcast channel $p_{Y_{[1:K]}|X}$ with side information Z_k at decoder k ,

$$\mathcal{R}_{\text{SI}}^{(i)} \subseteq \mathcal{R}_{\text{SI}} \subseteq \mathcal{R}_{\text{SI},1}^{(o)} \cap \mathcal{R}_{\text{SI},2}^{(o)} \subseteq \mathcal{R}_{\text{SI},1}'^{(o)}. \quad (70)$$

Remark 5. Similar to Theorem 1, Theorem 6 could also be extended to Gaussian or any other well-behaved continuous-alphabet source-channel pair, the problem of broadcasting Wyner-Ziv correlated sources, or Wyner-Ziv source broadcast with channel input cost.

B. Wyner-Ziv Source Broadcast over Degraded Channel with Degraded Side Information

Theorem 6 can be used to derive the inner bound and outer bound for the case of degraded channel and degraded side information. Define

$$\begin{aligned}
\mathcal{R}_{\text{SI-D}}^{(i)} = & \left\{ D_{[1:K]} : \text{There exists some pmf } p_{V_K|S, p_{V_{K-1}|V_K} \cdots p_{V_1|V_2}}, \right. \\
& \text{and functions } x(v_K, s), \hat{s}_k(v_k, y_k, z_k), 1 \leq k \leq K \text{ such that} \\
& \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, \\
& \left. I(S; V_k) \leq \sum_{j=1}^k I(Y_j Z_j; V_j | V_{j-1}), 1 \leq k \leq K, \text{ where } V_0 \triangleq \emptyset \right\}, \quad (71)
\end{aligned}$$

and where

$$\begin{aligned}
\mathcal{R}_{\text{SI-D}}^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmfs } p_{V_K|S} p_{V_{K-1}|V_K} \cdots p_{V_1|V_2}, p_X \right. \\
& \text{and functions } \hat{s}_k(v_k, z_k), 1 \leq k \leq K \text{ such that} \\
& \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, \\
& (I(V_k; U_k | U_{k-1} Z_k) : k \in [1:K]) \in \mathcal{B}_{DBC} \left(p_X p_{Y_{[1:K]}|X} \right) \\
& \text{for any pmf } p_{U_{K-1}|S} p_{U_{K-2}|U_{K-1}} \cdots p_{U_1|U_2}, U_0 \triangleq \emptyset, U_K \triangleq S, \\
& \left. \text{and } (I(V_k; S | Z_k) : k \in [1:K]) \in \mathcal{B}_{SRC} \left(p_X p_{Y_{[1:K]}|X} \right) \right\}, \tag{72}
\end{aligned}$$

where $\mathcal{B}_{DBC} \left(p_X p_{Y_{[1:K]}|X} \right)$ and $\mathcal{B}_{SRC} \left(p_X p_{Y_{[1:K]}|X} \right)$ are given in (31) and (32), respectively. Then we have the following theorem. The proof is analogous to that of Theorem 3, and is therefore omitted.

Theorem 7. For transmitting source S over degraded broadcast channel $p_{Y_{[1:K]}|X} (X \rightarrow Y_K \rightarrow Y_{K-1} \rightarrow \cdots \rightarrow Y_1)$ with degraded side information $Z_k (S \rightarrow Z_K \rightarrow Z_{K-1} \rightarrow \cdots \rightarrow Z_1)$ at decoder k ,

$$\mathcal{R}_{\text{SI-D}}^{(i)} \subseteq \mathcal{R}_{\text{SI}} \subseteq \mathcal{R}_{\text{SI-D}}^{(o)}. \tag{73}$$

Remark 6. $\mathcal{R}_{\text{SI-D}}^{(o)}$ can be also expressed as

$$\begin{aligned}
\mathcal{R}_{\text{SI-D}}^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmf } p_{V_K|S} p_{V_{K-1}|V_K} \cdots p_{V_1|V_2} \right. \\
& \text{and functions } \hat{s}_k(v_k, z_k), 1 \leq k \leq K \text{ such that} \\
& \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, \\
& \mathcal{B}_{DBC-SI} \left(p_S p_{Z_{[1:K]}|S} p_{V_{[1:K]}|S, Z_{[1:K]}} \right) \subseteq \mathcal{B}_{DBC} \left(p_X p_{Y_{[1:K]}|X} \right), \\
& \left. \mathcal{B}_{SRC-SI} \left(p_S p_{Z_{[1:K]}|S} p_{V_{[1:K]}|S, Z_{[1:K]}} \right) \supseteq \mathcal{B}_{SRC} \left(p_X p_{Y_{[1:K]}|X} \right) \right\}, \tag{74}
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{B}_{DBC-SI} \left(p_X p_{Z_{[1:K]}|X} p_{Y_{[1:K]}|X, Z_{[1:K]}} \right) = & \left\{ R_{[1:K]} : \text{There exists some pmf } p_{V_{K-1}|X} p_{V_{K-2}|V_{K-1}} \cdots p_{V_1|V_2} \text{ such that} \right. \\
& R_k \geq 0, \sum_{j=1}^k R_j \leq \sum_{j=1}^k I(Y_j; V_j | V_{j-1} Z_j), 1 \leq k \leq K, \text{ where } V_0 \triangleq \emptyset, V_K \triangleq X \left. \right\}, \tag{75}
\end{aligned}$$

and

$$\mathcal{B}_{SRC-SI} \left(p_X p_{Z_{[1:K]}|X} p_{Y_{[1:K]}|X, Z_{[1:K]}} \right) = \left\{ R_{[1:K]} : R_k \geq 0, \sum_{j=1}^k R_j \geq I(X; Y_k | Z_k), 1 \leq k \leq K \right\}. \tag{76}$$

C. Wyner-Ziv Gaussian Source Broadcast

Consider sending Gaussian source $S \sim \mathcal{N}(0, N_S)$ with quadratic distortion measure $d_k(s, \hat{s}) = d(s, \hat{s}) \triangleq (s - \hat{s})^2$ over power-constrained Gaussian broadcast channel $Y_k = X + W_k, 1 \leq k \leq K$ with $\mathbb{E}(X^2) \leq P$ and $W_k \sim$

$\mathcal{N}(0, N_k)$, $N_1 \geq N_2 \geq \dots \geq N_K$. Assume the side information Z_k observed by receiver k satisfies $S = Z_k + B_k$ with independent Gaussian variables $Z_k \sim \mathcal{N}(0, N_S - \beta_k)$ and $B_k \sim \mathcal{N}(0, \beta_k)$. Assume bandwidth mismatch factor is b . Then the inner bound of Theorem 7 could recover the existing results in the literature [17], [18], and the outer bound of Theorem 7 could be used to prove the following outer bound for Wyner-Ziv Gaussian source broadcast problem. The proof is given in Appendix E.

Theorem 8. *For transmitting Gaussian source S over Gaussian broadcast channel with degraded side information Z_k ($\beta_1 \geq \beta_2 \geq \dots \geq \beta_K$) at decoder k ,*

$$\mathcal{R}_{SI} \subseteq \mathcal{R}_{SI-D}^{(o)} \triangleq \left\{ D_{[1:K]} : \text{For any variables } +\infty = \tau_0 \geq \tau_1 \geq \dots \geq \tau_K = 0, \right. \\ \left. \frac{1}{b} \left(\frac{1}{2} \log \frac{(\beta_k + \tau_k)(D_k + \tau_{k-1})}{(D_k + \tau_k)(\beta_k + \tau_{k-1})} : k \in [1:K] \right) \in \mathcal{C}_{GBC} \right\}, \quad (77)$$

where \mathcal{C}_{GBC} denotes the capacity region of the Gaussian broadcast channel given in (57).

The bound of Theorem 8 is shown in Fig. 3.

D. Wyner-Ziv Binary Source Broadcast

Consider sending binary source $S \sim \text{Bern}(\frac{1}{2})$ with Hamming distortion measure $d_k(s, \hat{s}) = d(s, \hat{s}) \triangleq 0$, if $s = \hat{s}$; 1, otherwise, over binary broadcast channel $Y_k = X \oplus W_k$, $1 \leq k \leq K$ with $W_k \sim \text{Bern}(p_k)$, $\frac{1}{2} \geq p_1 \geq p_2 \geq \dots \geq p_K \geq 0$. Assume the side information Z_k observed by receiver k satisfies $S = Z_k \oplus B_k$ with independent variables $Z_k \sim \text{Bern}(\frac{1}{2})$ and $B_k \sim \text{Bern}(\beta_k)$. Assume bandwidth mismatch factor is b .

Let $V_1 = (U_1, X_1^b)$, $V_2 = (U_2, X^b)$, $V_3 = \emptyset$, and U_1 and U_2 are independent of X_1^b and X^b . S , U_2 and U_1 satisfy the distribution $p_{U_2|S}p_{U_1|U_2}$, where

$$p_{U_2|S} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{pmatrix} q_2 \bar{\alpha}_2 & q_2 \alpha_2 & \bar{q}_2 \\ q_2 \alpha_2 & q_2 \bar{\alpha}_2 & \bar{q}_2 \end{pmatrix} \end{matrix}, \quad (78)$$

$$p_{U_1|U_2} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \end{matrix} & \begin{pmatrix} q'_1 \bar{\alpha}'_1 & q'_1 \alpha'_1 & \bar{q}'_1 \\ q'_1 \alpha'_1 & q'_1 \bar{\alpha}'_1 & \bar{q}'_1 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix}. \quad (79)$$

with $0 \leq q_2, q'_1 \leq 1, 0 \leq \alpha_2, \alpha'_1 \leq \frac{1}{2}$. X^b and X_1^b satisfy $X_1^b = X^b \oplus B^b$, $X^b \sim b$ dimensional $\text{Bern}(\frac{1}{2})$, and $B^b \sim b$ dimensional $\text{Bern}(\theta)$ with $0 \leq \theta \leq \frac{1}{2}$. Denote $\alpha_1 = \alpha_2 \star \alpha'_1$, $q_1 = q_2 q'_1$, and set $x^b(v_2, s) = x^b$ and for $i = 1, 2$,

$$\hat{s}_i(v_i, y_i^b, z_i) = \begin{cases} z_i, & \text{if } \alpha_i \geq \beta_i \text{ or } \alpha_i < \beta_i, u_i = 2; \\ u_i, & \text{if } \alpha_i < \beta_i, u_i = 0, 1. \end{cases} \quad (80)$$

Substitute these random variables and functions into $\mathcal{R}_{SI}^{(i)}$ in Theorem 6, then we get the following performance (the hybrid coding reduces to a layered digital coding), which is tighter than that of the Layered Description Scheme (LDS) [17, Lem. 4].

Corollary 3 (Layered Digital Coding). *For transmitting binary source S with Hamming distortion measure over 2-user binary broadcast channel with side information Z_k at decoder k ,*

$$\begin{aligned} \mathcal{R}_{SI} \supseteq \mathcal{R}_{LDC}^{(i)} \triangleq & \left\{ (D_1, D_2) : 0 \leq q_1 \leq q_2 \leq 1, 0 \leq \alpha_2 \leq \alpha_1 \leq \frac{1}{2}, 0 \leq \theta \leq \frac{1}{2}, \right. \\ & q_1 r(\alpha_1, \beta_1) \leq b(1 - H_2(\theta \star p_1)), \\ & q_1 r(\alpha_1, \beta_2) \leq b(1 - H_2(\theta \star p_2)), \\ & q_2 r(\alpha_2, \beta_2) \leq b(1 - H_2(p_2)), \\ & q_1 r(\alpha_1, \beta_1) + (q_2 r(\alpha_2, \beta_2) - q_1 r(\alpha_1, \beta_2)) \leq b(1 - H_2(\theta \star p_1)) + b(H_2(\theta \star p_2) - H_2(p_2)), \\ & \left. D_i \leq q_i \min\{\alpha_i, \beta_i\} + (1 - q_i)\beta_i, i = 1, 2 \right\}, \end{aligned} \quad (81)$$

where

$$r(\alpha, \beta) = H_2(\alpha \star \beta) - H_2(\alpha), \quad (82)$$

\star denotes the binary convolution given in (59), and H_2 denotes the binary entropy function given in (60).

In addition, the outer bound of Theorem 7 reduces to the following one for Wyner-Ziv binary case. The proof is given in Appendix F.

Theorem 9. *For transmitting binary source S with Hamming distortion measure over K -user binary broadcast channel with degraded side information Z_k ($\frac{1}{2} \geq \beta_1 \geq \beta_2 \geq \dots \geq \beta_K \geq 0$) at decoder k ,*

$$\begin{aligned} \mathcal{R}_{SI} \subseteq \mathcal{R}_{SI-D}^{(o)} \triangleq & \left\{ D_{[1:K]} : \text{There exists some variables } 0 \leq \alpha_1, \alpha_2, \dots, \alpha_K \leq \frac{1}{2} \right. \\ & \text{such that } \alpha_k \leq D'_k \triangleq \min\{D_k, \beta_k\}, 1 \leq k \leq K, \\ & \text{and for any variables } \frac{1}{2} = \tau_0 \geq \tau_1 \geq \tau_2 \geq \dots \geq \tau_K = 0, \\ & \left. \frac{1}{b} \left(\eta_k (H_2(\beta_k \star \tau_k) - H_2(\beta_k \star \tau_{k-1}) - (H_4(\alpha_k, \beta_k, \tau_k) - H_4(\alpha_k, \beta_k, \tau_{k-1}))) : k \in [1:K] \right) \in \mathcal{C}_{BBC} \right\}, \end{aligned} \quad (83)$$

where \mathcal{C}_{BBC} denotes the capacity region of the binary broadcast channel given in (62),

$$\eta_k \triangleq \begin{cases} \frac{\beta_k - D'_k}{\beta_k - \alpha_k}, & \text{if } \alpha_k < \beta_k, \\ 0, & \text{otherwise,} \end{cases} \quad (84)$$

$$\begin{aligned} H_4(x, y, z) \triangleq & -(xyz + \overline{xy}\overline{z}) \log(xyz + \overline{xy}\overline{z}) - (x\overline{y}z + \overline{x}y\overline{z}) \log(x\overline{y}z + \overline{x}y\overline{z}) \\ & - (xy\overline{z} + \overline{x}\overline{y}z) \log(xy\overline{z} + \overline{x}\overline{y}z) - (x\overline{y}\overline{z} + \overline{x}yz) \log(x\overline{y}\overline{z} + \overline{x}yz), \end{aligned} \quad (85)$$

and $\bar{x} \triangleq 1 - x$.

The bounds in Corollary 3 and Theorem 9 are shown in Fig. 4.

V. CONCLUDING REMARKS

In this paper, we focused on the joint source-channel coding problem of sending a memoryless source over memoryless broadcast channel, and developed an inner bound and several outer bounds for this problem. The inner bound is achieved by a unified hybrid coding scheme, and it can recover the best known performance of existing hybrid coding. Similarly, our outer bounds can also recover the best known outer bound in the literature. Besides, we also extend the results to Wyner-Ziv source broadcast problem. All these bounds are also used to generate some new results, including the bounds on capacity region of general broadcast channel with common messages which respectively generalize Marton's inner bound and Nair-El Gamal outer bound.

The inner bound achieved by proposed hybrid coding is established by using generalized Multivariate Covering Lemma and generalized Multivariate Packing Lemma, and the outer bounds are derived by introducing auxiliary random variables (at sender side or receiver sides) and exploiting Csiszár sum identity as in [12]. These lemmas and tools are expected to be exploited to derive more and stronger achievability and converse results for network information theory.

APPENDIX A

PROOF OF LEMMA 1

We follow similar steps to the proof of mutual covering lemma [15]. Let

$$\mathcal{B} = \left\{ m_{[1:k]} \in \prod_{i=1}^k [1 : 2^{nr_i}] : (U^n, V_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_\epsilon^{(n)} \right\}. \quad (86)$$

Then we only need to show $\lim_{n \rightarrow \infty} \mathbb{P}(|\mathcal{B}| = 0) = 0$. On the other hand,

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\mathcal{B}| = 0) = \lim_{n \rightarrow \infty} \sum_{u^n, v_0^n} p_{U^n, V_0^n}(u^n, v_0^n) \mathbb{P}(|\mathcal{B}| = 0 | u^n, v_0^n) \quad (87)$$

$$\leq \lim_{n \rightarrow \infty} \sum_{(u^n, v_0^n) \in \mathcal{T}_{\epsilon'}^{(n)}} p_{U^n, V_0^n}(u^n, v_0^n) \mathbb{P}(|\mathcal{B}| = 0 | u^n, v_0^n) + \lim_{n \rightarrow \infty} \mathbb{P}\left((u^n, v_0^n) \notin \mathcal{T}_{\epsilon'}^{(n)}\right) \quad (88)$$

$$= \lim_{n \rightarrow \infty} \sum_{(u^n, v_0^n) \in \mathcal{T}_{\epsilon'}^{(n)}} p_{U^n, V_0^n}(u^n, v_0^n) \mathbb{P}(|\mathcal{B}| = 0 | u^n, v_0^n) \quad (89)$$

To prove $\lim_{n \rightarrow \infty} \mathbb{P}(|\mathcal{B}| = 0) = 0$, it is sufficient to show $\lim_{n \rightarrow \infty} \mathbb{P}(|\mathcal{B}| = 0 | u^n, v_0^n) = 0$ for any $(u^n, v_0^n) \in \mathcal{T}_{\epsilon'}^{(n)}$.

Utilizing Chebyshev lemma [16, App. B], we can bound the probability as

$$\mathbb{P}(|\mathcal{B}| = 0 | u^n, v_0^n) \leq \mathbb{P}((|\mathcal{B}| - \mathbb{E}|\mathcal{B}|)^2 \geq (E|\mathcal{B}|)^2 | u^n, v_0^n) \leq \frac{\text{Var}(|\mathcal{B}| | u^n, v_0^n)}{(\mathbb{E}|\mathcal{B}| | u^n, v_0^n)^2}. \quad (90)$$

Next we prove the upper bound $\frac{\text{Var}(|\mathcal{B}||u^n, v_0^n)}{(\mathbb{E}(|\mathcal{B}||u^n, v_0^n))^2}$ tends to zero as $n \rightarrow \infty$. Define

$$E(m_{[1:k]}) \triangleq \begin{cases} 1, & \text{if } (u^n, v_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_\epsilon^{(n)}; \\ 0, & \text{otherwise,} \end{cases} \quad (91)$$

for each $m_{[1:k]} \in \prod_{i=1}^k [1 : 2^{nr_i}]$, then $|\mathcal{B}|$ can be expressed as

$$|\mathcal{B}| = \sum_{m_{[1:k]} \in \prod_{i=1}^k [1 : 2^{nr_i}]} E(m_{[1:k]}). \quad (92)$$

Denote

$$p_0 = \mathbb{P}\left((u^n, v_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_\epsilon^{(n)} | u^n, v_0^n\right), \quad (93)$$

$$p_{\mathcal{I}} = \mathbb{P}\left((u^n, v_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_\epsilon^{(n)}, (u^n, v_0^n, V_{[1:k]}^n(m_{\mathcal{I}}, m'_{\mathcal{I}^c})) \in \mathcal{T}_\epsilon^{(n)} | u^n, v_0^n\right), \quad (94)$$

for $m_{[1:k]} = \mathbf{1}$, and $m'_{[1:k]} = \mathbf{2}$. Obviously, $p_{[1:k]} = p_0$. Then

$$\mathbb{E}(|\mathcal{B}||u^n, v_0^n) = \sum_{m_{[1:k]}} \mathbb{P}\left((u^n, v_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_\epsilon^{(n)} | u^n, v_0^n\right) = 2^{n \sum_{j=1}^k r_j} p_0, \quad (95)$$

and

$$\begin{aligned} & \mathbb{E}(|\mathcal{B}|^2 | u^n, v_0^n) \\ &= \sum_{\mathcal{I} \subseteq [1:k]} \sum_{m_{[1:k]}} \sum_{m'_{\mathcal{I}^c} : m'_{\mathcal{I}^c} \not\Leftarrow m_{\mathcal{I}^c}} \mathbb{P}\left((u^n, v_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_\epsilon^{(n)}, (u^n, v_0^n, V_{[1:k]}^n(m_{\mathcal{I}}, m'_{\mathcal{I}^c})) \in \mathcal{T}_\epsilon^{(n)} | u^n, v_0^n\right) \\ &= 2^{n \sum_{j=1}^k r_j} p_0 + \sum_{\mathcal{I} \subsetneq [1:k]} \sum_{m_{[1:k]}} \sum_{m'_{\mathcal{I}^c} : m'_{\mathcal{I}^c} \not\Leftarrow m_{\mathcal{I}^c}} \mathbb{P}\left((u^n, v_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_\epsilon^{(n)}, (u^n, v_0^n, V_{[1:k]}^n(m_{\mathcal{I}}, m'_{\mathcal{I}^c})) \in \mathcal{T}_\epsilon^{(n)} | u^n, v_0^n\right). \end{aligned} \quad (96)$$

(97)

Define

$$\mathbb{J} \triangleq \{\mathcal{J} \subsetneq [1:k] : \text{if } j \in \mathcal{J}, \text{ then } \mathcal{A}_j \subseteq \mathcal{J}\}. \quad (98)$$

Then any set $\mathcal{I} \subsetneq [1:k]$ can transform into a $\mathcal{J}(\mathcal{I}) \in \mathbb{J}$ by removing all the elements j 's such that $\mathcal{A}_j \not\subseteq \mathcal{I}$.

According to generation of random codebook, we can observe that $p_{\mathcal{I}} = p_{\mathcal{J}(\mathcal{I})}$. Therefore,

$$\begin{aligned} & \sum_{\mathcal{I} \subsetneq [1:k]} \sum_{m_{[1:k]}} \sum_{m'_{\mathcal{I}^c} : m'_{\mathcal{I}^c} \not\Leftarrow m_{\mathcal{I}^c}} \mathbb{P}\left((u^n, v_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_\epsilon^{(n)}, (u^n, v_0^n, V_{[1:k]}^n(m_{\mathcal{I}}, m'_{\mathcal{I}^c})) \in \mathcal{T}_\epsilon^{(n)} | u^n, v_0^n\right) \\ &= \sum_{\mathcal{I} \subsetneq [1:k]} \sum_{m_{[1:k]}} \sum_{m'_{\mathcal{I}^c} : m'_{\mathcal{I}^c} \not\Leftarrow m_{\mathcal{I}^c}} p_{\mathcal{J}(\mathcal{I})} \end{aligned} \quad (99)$$

$$\leq \sum_{\mathcal{I} \subsetneq [1:k]} 2^{n(\sum_{j=1}^k r_j + \sum_{j \in \mathcal{I}^c} r_j)} p_{\mathcal{J}(\mathcal{I})} \quad (100)$$

$$\leq \sum_{\mathcal{I} \subsetneq [1:k]} 2^{n(\sum_{j=1}^k r_j + \sum_{j \in (\mathcal{J}(\mathcal{I}))^c} r_j)} p_{\mathcal{J}(\mathcal{I})} \quad (101)$$

$$\leq \sum_{\mathcal{J} \in \mathbb{J}} 2^{k-|\mathcal{J}|} 2^{n(\sum_{j=1}^k r_j + \sum_{j \in \mathcal{J}^c} r_j)} p_{\mathcal{J}} \quad (102)$$

$$\leq \sum_{\mathcal{J} \in \mathbb{J}} 2^{n(\sum_{j=1}^k r_j + \sum_{j \in \mathcal{J}^c} r_j + o(1))} p_{\mathcal{J}}, \quad (103)$$

where (101) follows from $\mathcal{J}(\mathcal{I}) \subseteq \mathcal{I}$, (102) follows from that for each $\mathcal{J} \subseteq \mathbb{J}$, there are at most $2^{k-|\mathcal{J}|}$ of \mathcal{I} 's that could transform into \mathcal{J} , and $o(1)$ denotes a term that vanishes as $n \rightarrow \infty$. Hence

$$\text{Var}(|\mathcal{B}||u^n, v_0^n) \leq \mathbb{E}(|\mathcal{B}|^2|u^n, v_0^n) \quad (104)$$

$$\leq 2^{n \sum_{j=1}^k r_j} p_0 + \sum_{\mathcal{J} \in \mathbb{J}} 2^{n(\sum_{j=1}^k r_j + \sum_{j \in \mathcal{J}^c} r_j + o(1))} p_{\mathcal{J}}. \quad (105)$$

Furthermore we have

$$\frac{\text{Var}(|\mathcal{B}||u^n, v_0^n)}{(\mathbb{E}(|\mathcal{B}||u^n, v_0^n))^2} \leq \frac{2^{n \sum_{j=1}^k r_j} p_0 + \sum_{\mathcal{J} \in \mathbb{J}} 2^{n(\sum_{j=1}^k r_j + \sum_{j \in \mathcal{J}^c} r_j + o(1))} p_{\mathcal{J}}}{\left(2^{n \sum_{j=1}^k r_j} p_0\right)^2} \quad (106)$$

$$= 2^{-n \sum_{j=1}^k r_j} \frac{1}{p_0} + \sum_{\mathcal{J} \in \mathbb{J}} 2^{n(-\sum_{j \in \mathcal{J}} r_j + o(1))} \frac{p_{\mathcal{J}}}{p_0^2}. \quad (107)$$

According to generation process of random codebook, we can observe that

$$p_0 = \sum_{v_{[1:k]}^n : (u^n, v_0^n, v_{[1:k]}^n) \in \mathcal{T}_\epsilon^{(n)}} \mathbb{P}\left(V_{[1:k]}^n(m_{[1:k]}) = v_{[1:k]}^n | u^n, v_0^n\right) \quad (108)$$

$$\geq 2^{-n(\sum_{j=1}^k H(V_j|V_{\mathcal{A}_j} V_0) - H(V_{[1:k]}|UV_0) + 2\delta(\epsilon))}, \quad (109)$$

where (109) follows from that for any $(u^n, v_0^n, v_{[1:k]}^n) \in \mathcal{T}_\epsilon^{(n)}$,

$$\mathbb{P}\left(V_{[1:k]}^n(m_{[1:k]}) = v_{[1:k]}^n | u^n, v_0^n\right) \geq 2^{-n(\sum_{j=1}^k H(V_j|V_{\mathcal{A}_j} V_0) + \delta(\epsilon))}, \quad (110)$$

and for any $(u^n, v_0^n) \in \mathcal{T}_\epsilon^{(n)}$,

$$\left|\left\{v_{[1:k]}^n : (u^n, v_0^n, v_{[1:k]}^n) \in \mathcal{T}_\epsilon^{(n)}\right\}\right| \geq 2^{n(H(V_{[1:k]}|UV_0) + \delta(\epsilon))}. \quad (111)$$

Similarly, we also can get

$$p_{\mathcal{J}} \leq 2^{-n(\sum_{j=1}^k H(V_j|V_{\mathcal{A}_j} V_0) + \sum_{j \in \mathcal{J}^c} H(V_j|V_{\mathcal{A}_j} V_0) - H(V_{[1:k]}|UV_0) - H(V_{\mathcal{J}^c}|UV_0 V_{\mathcal{J}}) - 4\delta(\epsilon))}. \quad (112)$$

Substitute (109) and (112) into (107), then we have

$$\begin{aligned} \frac{\text{Var}(|\mathcal{B}||u^n, v_0^n)}{(\mathbb{E}(|\mathcal{B}||u^n, v_0^n))^2} &\leq 2^{-n(\sum_{j=1}^k r_j - (\sum_{j=1}^k H(V_j|V_{\mathcal{A}_j} V_0) - H(V_{[1:k]}|UV_0) + 2\delta(\epsilon)))} \\ &\quad + \sum_{\mathcal{J} \in \mathbb{J}} 2^{-n(\sum_{j \in \mathcal{J}} r_j - (\sum_{j \in \mathcal{J}} H(V_j|V_{\mathcal{A}_j} V_0) - H(V_{\mathcal{J}}|UV_0) + 6\delta(\epsilon) + o(1)))}. \end{aligned} \quad (113)$$

(113) tends to zero if

$$\sum_{j=1}^k r_j > \sum_{j=1}^k H(V_j|V_{\mathcal{A}_j} V_0) - H(V_{[1:k]}|UV_0) + 2\delta(\epsilon) \quad (114)$$

$$\sum_{j \in \mathcal{J}} r_j > \sum_{j \in \mathcal{J}} H(V_j|V_{\mathcal{A}_j} V_0) - H(V_{\mathcal{J}}|UV_0) + 6\delta(\epsilon) + o(1), \quad (115)$$

i.e., $\sum_{j \in \mathcal{J}} r_j > \sum_{j \in \mathcal{J}} H(V_j|V_{\mathcal{A}_j} V_0) - H(V_{\mathcal{J}}|UV_0) + \delta'(\epsilon)$ for some $\delta'(\epsilon)$ that tends to zero as $\epsilon \rightarrow 0$. This completes the proof.

APPENDIX B
PROOF OF LEMMA 2

For any \mathcal{J} such that $\mathcal{J} \neq \emptyset$ and if $j \in \mathcal{J}$ then $\mathcal{A}_j \subseteq \mathcal{J}$,

$$\begin{aligned} & \mathbb{P} \left((U^n, V_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m_{[1:k]} \right) \\ & \leq \mathbb{P} \left((U^n, V_0^n, V_{\mathcal{J}}^n(m_{\mathcal{J}})) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m_{\mathcal{J}} \right) \end{aligned} \quad (116)$$

$$= \sum_{u^n, v_0^n} p_{U^n, V_0^n}(u^n, v_0^n) \mathbb{P} \left((u^n, v_0^n, V_{\mathcal{J}}^n(m_{\mathcal{J}})) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m_{\mathcal{J}} | u^n, v_0^n \right) \quad (117)$$

$$\leq \sum_{u^n, v_0^n} p_{U^n, V_0^n}(u^n, v_0^n) \sum_{m_{\mathcal{J}}} \mathbb{P} \left((u^n, v_0^n, V_{\mathcal{J}}^n(m_{\mathcal{J}})) \in \mathcal{T}_\epsilon^{(n)} | u^n, v_0^n \right). \quad (118)$$

Similar to (109), we can obtain that

$$\mathbb{P} \left((u^n, v_0^n, V_{\mathcal{J}}^n(m_{\mathcal{J}})) \in \mathcal{T}_\epsilon^{(n)} | u^n, v_0^n \right) \leq 2^{-n(\sum_{j \in \mathcal{J}} H(V_j | V_{\mathcal{A}_j} V_0) - H(V_{\mathcal{J}} | UV_0) - 2\delta(\epsilon))}. \quad (119)$$

Substitute it into (118), then we have

$$\begin{aligned} & \mathbb{P} \left((U^n, V_0^n, V_{[1:k]}^n(m_{[1:k]})) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m_{[1:k]} \right) \\ & \leq 2^{n(\sum_{j \in \mathcal{J}} r_j - (\sum_{j \in \mathcal{J}} H(V_j | V_{\mathcal{A}_j} V_0) - H(V_{\mathcal{J}} | UV_0) - 2\delta(\epsilon)))}. \end{aligned} \quad (120)$$

(120) tends to zero if

$$\sum_{j \in \mathcal{J}} r_j < \sum_{j \in \mathcal{J}} H(V_j | V_{\mathcal{A}_j} V_0) - H(V_{\mathcal{J}} | UV_0) - 2\delta(\epsilon). \quad (121)$$

This completes the proof.

APPENDIX C
PROOF OF THEOREM 1

A. Inner Bound

Actually the inner bound can be seen as a corollary to [20, Thm. 1] by choosing proper network topology, transit probability and symbol-by-symbol functions. For completeness and clarity, next we provide a direct description of the proposed hybrid coding scheme and a direct proof for it.

Codebook Generation: Fix conditional pmf $p_{V_{[1:N]}|S}$, vector $r_{[1:N]}$, encoding function $x(v_{[1:N]}, s)$ and decoding

functions $\hat{s}_k(v_{\mathcal{D}_k}, y_k)$ that satisfy

$$\mathbb{E}d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \quad (122)$$

$$\sum_{j \in \mathcal{J}} r_j > \sum_{j \in \mathcal{J}} H(V_j | V_{\mathcal{A}_j}) - H(V_{\mathcal{J}} | S) \quad (123)$$

for all $\mathcal{J} \subseteq [1 : N]$ such that $\mathcal{J} \neq \emptyset$ and if $j \in \mathcal{J}$, then $\mathcal{A}_j \subseteq \mathcal{J}$,

$$\sum_{j \in \mathcal{J}^c} r_j < \sum_{j \in \mathcal{J}^c} H(V_j | V_{\mathcal{A}_j}) - H(V_{\mathcal{J}^c} | Y_k V_{\mathcal{J}}) \quad (124)$$

for all $1 \leq k \leq K$ and for all $\mathcal{J} \subseteq \mathcal{D}_k$ such that $\mathcal{J}^c \triangleq \mathcal{D}_k \setminus \mathcal{J} \neq \emptyset$ and if $j \in \mathcal{J}$, then $\mathcal{A}_j \subseteq \mathcal{J}$.

For each $j \in [1 : N]$ and each $m_{\mathcal{A}_j} \in \prod_{i \in \mathcal{A}_j} [1 : 2^{nr_i}]$, randomly and independently generate a set of sequences $v_j^n(m_{\mathcal{A}_j}, m_j)$, $m_j \in [1 : 2^{nr_j}]$, with each distributed according to $\prod_{i=1}^n p_{V_j | V_{\mathcal{A}_j}}(v_{j,i} | v_{\mathcal{A}_j,i}(m_{\mathcal{A}_j}))$. The codebook

$$\mathcal{C} = \left\{ v_{[1:N]}^n(m_{[1:N]}) : m_{[1:N]} \in \prod_{i=1}^N [1 : 2^{nr_i}] \right\}. \quad (125)$$

is revealed to the encoder and all the decoders.

Encoding: We use joint typicality encoding. Given s^n , encoder finds the smallest index vector $m_{[1:N]}$ such that $(s^n, v_{[1:N]}^n(m_{[1:N]})) \in \mathcal{T}_{\epsilon}^{(n)}$. If there is no such index vector, let $m_{[1:N]} = \mathbf{1}$. Then the encoder transmits the signal

$$x_i = x(v_{[1:N],i}(m_{[1:N]}), s_i), 1 \leq i \leq n. \quad (126)$$

Decoding: We use joint typicality decoding. Let $\epsilon' > \epsilon$. Upon receiving signal y_k^n , the decoder of the receiver k finds the smallest index vector $\hat{m}_{\mathcal{D}_k}^{(k)}$ such that

$$(v_{\mathcal{D}_k}^n(\hat{m}_{\mathcal{D}_k}^{(k)}), y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)}. \quad (127)$$

If there is no such index vector, let $\hat{m}_{\mathcal{D}_k}^{(k)} = \mathbf{1}$. The decoder reconstructs the source as

$$\hat{s}_{k,i} = \hat{s}_k(v_{\mathcal{D}_k,i}(\hat{m}_{\mathcal{D}_k}^{(k)}), y_{k,i}), 1 \leq i \leq n. \quad (128)$$

Analysis of Expected Distortion: We bound the distortion averaged over S^n , and the random choice of the codebook \mathcal{C} . Define the “error” event

$$\mathcal{E} = \mathcal{E}_1 \cup \left(\bigcup_k \mathcal{E}_{2,k} \right) \cup \left(\bigcup_k \mathcal{E}_{3,k} \right), \quad (129)$$

where

$$\mathcal{E}_1 = \left\{ (S^n, V_{[1:N]}^n(m_{[1:N]})) \notin \mathcal{T}_{\epsilon}^{(n)} \text{ for all } m_{[1:N]} \right\}, \quad (130)$$

$$\mathcal{E}_{2,k} = \left\{ (S^n, V_{[1:N]}^n(M_{[1:N]}), Y_k^n) \notin \mathcal{T}_{\epsilon'}^{(n)} \right\}, \quad (131)$$

$$\mathcal{E}_{3,k} = \left\{ (V_{\mathcal{D}_k}^n(m'_{\mathcal{D}_k}), Y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \text{ for some } m'_{\mathcal{D}_k}, m'_{\mathcal{D}_k} \neq M_{\mathcal{D}_k} \right\}, \quad (132)$$

for $1 \leq k \leq K$. Using union bound, we have

$$\mathbb{P}(\mathcal{E}) \leq \mathbb{P}(\mathcal{E}_1) + \sum_{k=1}^K \mathbb{P}(\mathcal{E}_1^c \cap \mathcal{E}_{2,k}) + \sum_{k=1}^K \mathbb{P}(\mathcal{E}_{3,k}). \quad (133)$$

Now we claim that if (123) and (124) hold, then $\mathbb{P}(\mathcal{E})$ tends to zero as $n \rightarrow \infty$. Before proving it, we show that this claim implies the inner bound of Theorem 1.

Define

$$\mathcal{E}_{4,k} = \left\{ (S^n, V_{\mathcal{D}_k}^n(\hat{M}_{\mathcal{D}_k}^{(k)}), Y_k^n) \notin \mathcal{T}_{\epsilon'}^{(n)} \right\}. \quad (134)$$

then we have $\mathcal{E}^c \subseteq \mathcal{E}_{4,k}^c$, i.e., $\mathcal{E}_{4,k} \subseteq \mathcal{E}$. This implies that $\mathbb{P}(\mathcal{E}_{4,k}) \leq \mathbb{P}(\mathcal{E}) \rightarrow 0$ as $n \rightarrow \infty$. Then utilizing typical average lemma [16], we have

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \mathbb{E} d_k(S^n, \hat{S}_k^n) \\ &= \limsup_{n \rightarrow \infty} \left(\mathbb{P}(\mathcal{E}_{4,k}) \mathbb{E} \left[d_k(S^n, \hat{S}_k^n) | \mathcal{E}_{4,k} \right] + \mathbb{P}(\mathcal{E}_{4,k}^c) \mathbb{E} \left[d_k(S^n, \hat{S}_k^n) | \mathcal{E}_{4,k}^c \right] \right) \end{aligned} \quad (135)$$

$$= \limsup_{n \rightarrow \infty} \mathbb{E} \left[d_k(S^n, \hat{S}_k^n) | \mathcal{E}_{4,k}^c \right] \quad (136)$$

$$\leq (1 + \epsilon') \mathbb{E} d_k(S, \hat{S}_k) \quad (137)$$

$$\leq (1 + \epsilon') D_k. \quad (138)$$

Therefore, the desired distortions are achieved for sufficiently small ϵ' .

Next we turn back to prove the claim above. Following from Multivariate Covering Lemma (Lemma 1), the first term of (133), $\mathbb{P}(\mathcal{E}_1)$, vanishes as $n \rightarrow \infty$, and according to conditional typicality lemma [16, Sec. 3.7], the second item tends to zero as $n \rightarrow \infty$.

Now we focus on the third term of (133). $\mathcal{E}_{3,k}$ can be written as

$$\mathcal{E}_{3,k} = \bigcup_{\mathcal{I} \subseteq \mathcal{D}_k} \mathcal{E}_{3,k}^{\mathcal{I}}, \quad (139)$$

where

$$\mathcal{E}_{3,k}^{\mathcal{I}} = \left\{ (V_{\mathcal{D}_k}^n(M_{\mathcal{I}}, m'_{\mathcal{I}^c}), Y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \text{ for some } m'_{\mathcal{I}^c}, m'_{\mathcal{I}^c} \not\Leftarrow M_{\mathcal{I}^c} \right\}, \quad (140)$$

with $\mathcal{I}^c \triangleq \mathcal{D}_k \setminus \mathcal{I}$. Using union bound we have

$$\mathbb{P}(\mathcal{E}_{3,k}) \leq \sum_{\mathcal{I} \subseteq \mathcal{D}_k} \mathbb{P}(\mathcal{E}_{3,k}^{\mathcal{I}}). \quad (141)$$

Each \mathcal{D}_k has finite number of subsets, hence we only need to show for each $\mathcal{I} \subseteq \mathcal{D}_k$, $\mathbb{P}(\mathcal{E}_{3,k}^{\mathcal{I}})$ vanishes as $n \rightarrow \infty$. To show this, it is needed to analyze the correlation between coding index $M_{[1:N]}$ and nonchosen codewords. Specifically, $M_{[1:N]}$ depends on source sequence and the entire codebook, and hence standard packing lemma cannot be applied directly. This problem has been resolved by the technique developed in [15], [20].

$$\begin{aligned} & \mathbb{P} \left((V_{\mathcal{D}_k}^n (M_{\mathcal{I}}, m'_{\mathcal{I}^c}), Y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \text{ for some } m'_{\mathcal{I}^c}, m'_{\mathcal{I}^c} \not\Leftarrow M_{\mathcal{I}^c} \right) \\ &= \sum_{m_{[1:N]}} \sum_{y_k^n} \mathbb{P} \left(M_{[1:N]} = m_{[1:N]}, Y_k^n = y_k^n, (V_{\mathcal{D}_k}^n (m_{\mathcal{I}}, m'_{\mathcal{I}^c}), y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \text{ for some } m'_{\mathcal{I}^c}, m'_{\mathcal{I}^c} \not\Leftarrow m_{\mathcal{I}^c} \right) \end{aligned} \quad (142)$$

$$\leq \sum_{m_{[1:N]}} \sum_{y_k^n} \sum_{m'_{\mathcal{I}^c}: m'_{\mathcal{I}^c} \not\Leftarrow m_{\mathcal{I}^c}} \mathbb{P} \left(M_{[1:N]} = m_{[1:N]}, Y_k^n = y_k^n, (V_{\mathcal{D}_k}^n (m_{\mathcal{I}}, m'_{\mathcal{I}^c}), y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \right) \quad (143)$$

where (143) follows from the union bound.

Define a sub-codebook as

$$\mathcal{C}_{(m_{\mathcal{I}}, m'_{\mathcal{I}^c})} = \left\{ V_{[1:N]}^n \left(m_{\mathcal{I}}, m''_{\mathcal{I}^c}, m''_{\mathcal{D}_k^c} \right) : \forall \left(m''_{\mathcal{I}^c}, m''_{\mathcal{D}_k^c} \right), m''_{\mathcal{I}^c} \not\Leftarrow m'_{\mathcal{I}^c} \right\}. \quad (144)$$

Define another coding index as $\tilde{M}_{[1:N]}$ which is generated by performing the same coding process as $M_{[1:N]}$ but on codebook $\mathcal{C}_{(m_{\mathcal{I}}, m'_{\mathcal{I}^c})}$, i.e., given source sequence s^n , encoder finds the smallest index vector $\tilde{m}_{[1:N]}$ such that $(s^n, v_{[1:N]}^n(\tilde{m}_{[1:N]})) \in \mathcal{T}_{\epsilon'}^{(n)}$; if there is no such index vector, let $\tilde{m}_{[1:N]} = \mathbf{1}$. Then according to the generation process of $M_{[1:N]}$ and $\tilde{M}_{[1:N]}$, we have if $M_{[1:N]} = m_{[1:N]}$, then $\tilde{M}_{[1:N]} = m_{[1:N]}$. Now continuing with (143), we have

$$\begin{aligned} & \mathbb{P} \left(M_{[1:N]} = m_{[1:N]}, Y_k^n = y_k^n, (V_{\mathcal{D}_k}^n (m_{\mathcal{I}}, m'_{\mathcal{I}^c}), y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \right) \\ &= \sum_{v_{\mathcal{D}_k}^n, c, s^n} \mathbb{P} \left(M_{[1:N]} = m_{[1:N]}, \mathcal{C}_{(m_{\mathcal{I}}, m'_{\mathcal{I}^c})} = c, S^n = s^n, V_{\mathcal{D}_k}^n (m_{\mathcal{I}}, m'_{\mathcal{I}^c}) = v_{\mathcal{D}_k}^n \right) \\ & \quad \prod_{i=1}^n p_{Y_k|X} (y_{k,i} | x(v_{[1:N],i}(m_{[1:N]}), s_i)) \mathbf{1} \left\{ (v_{\mathcal{D}_k}^n, y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \right\} \end{aligned} \quad (145)$$

$$\begin{aligned} &= \sum_{v_{\mathcal{D}_k}^n, c, s^n} \mathbb{P} \left(M_{[1:N]} = m_{[1:N]}, \tilde{M}_{[1:N]} = m_{[1:N]}, \mathcal{C}_{(m_{\mathcal{I}}, m'_{\mathcal{I}^c})} = c, S^n = s^n, V_{\mathcal{D}_k}^n (m_{\mathcal{I}}, m'_{\mathcal{I}^c}) = v_{\mathcal{D}_k}^n \right) \\ & \quad \prod_{i=1}^n p_{Y_k|X} (y_{k,i} | x(v_{[1:N],i}(m_{[1:N]}), s_i)) \mathbf{1} \left\{ (v_{\mathcal{D}_k}^n, y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \right\} \end{aligned} \quad (146)$$

$$\begin{aligned} &\leq \sum_{v_{\mathcal{D}_k}^n, c, s^n} \mathbb{P} \left(\tilde{M}_{[1:N]} = m_{[1:N]}, \mathcal{C}_{(m_{\mathcal{I}}, m'_{\mathcal{I}^c})} = c, S^n = s^n, V_{\mathcal{D}_k}^n (m_{\mathcal{I}}, m'_{\mathcal{I}^c}) = v_{\mathcal{D}_k}^n \right) \\ & \quad \prod_{i=1}^n p_{Y_k|X} (y_{k,i} | x(v_{[1:N],i}(m_{[1:N]}), s_i)) \mathbf{1} \left\{ (v_{\mathcal{D}_k}^n, y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \right\} \end{aligned} \quad (147)$$

$$\begin{aligned} &= \sum_{v_{\mathcal{D}_k}^n, c, s^n} \mathbb{P} \left(\tilde{M}_{[1:N]} = m_{[1:N]}, \mathcal{C}_{(m_{\mathcal{I}}, m'_{\mathcal{I}^c})} = c, S^n = s^n \right) \prod_{i=1}^n p_{Y_k|X} (y_{k,i} | x(v_{[1:N],i}(m_{[1:N]}), s_i)) \\ & \quad \mathbb{P} \left(V_{\mathcal{D}_k}^n (m_{\mathcal{I}}, m'_{\mathcal{I}^c}) = v_{\mathcal{D}_k}^n | \mathcal{C}_{(m_{\mathcal{I}}, m'_{\mathcal{I}^c})} = c \right) \mathbf{1} \left\{ (v_{\mathcal{D}_k}^n, y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \right\} \end{aligned} \quad (148)$$

where $c = \left\{ v_{[1:N]}^n (m_{\mathcal{I}}, m''_{\mathcal{I}^c}, m''_{\mathcal{D}_k^c}) : \forall (m''_{\mathcal{I}^c}, m''_{\mathcal{D}_k^c}), m''_{\mathcal{I}^c} \not\Leftarrow m'_{\mathcal{I}^c} \right\}$, and (148) follows from the fact that $V_{\mathcal{D}_k}^n (m_{\mathcal{I}}, m'_{\mathcal{I}^c}) \rightarrow \mathcal{C}_{(m_{\mathcal{I}}, m'_{\mathcal{I}^c})} \rightarrow (S^n, \tilde{M}_{[1:N]})$ forms a Markov chain.

Define

$$\mathbb{J} \triangleq \{ \mathcal{J} \subseteq \mathcal{D}_k : \text{if } j \in \mathcal{J}, \text{ then } \mathcal{A}_j \subseteq \mathcal{J} \}. \quad (149)$$

Then any set $\mathcal{I} \subseteq \mathcal{D}_k$ can transform into a $\mathcal{J}(\mathcal{I}) \in \mathbb{J}$ by removing all the elements j 's such that $\mathcal{A}_j \not\subseteq \mathcal{I}$. Denote $\mathcal{J}^c \triangleq \mathcal{D}_k \setminus \mathcal{J}$. Then according to the generation process of the codebook, continuing with (148), we have

$$\begin{aligned} & \sum_{v_{\mathcal{D}_k}^n} \mathbb{P} \left(V_{\mathcal{D}_k}^n(m_{\mathcal{I}}, m'_{\mathcal{I}^c}) = v_{\mathcal{D}_k}^n | \mathcal{C}_{(m_{\mathcal{I}}, m'_{\mathcal{I}^c})} = c \right) 1 \left\{ (v_{\mathcal{D}_k}^n, y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \right\} \\ &= \sum_{v_{\mathcal{J}^c}^n} \mathbb{P} \left(V_{\mathcal{J}^c}^n(m_{\mathcal{I}}, m'_{\mathcal{I}^c}) = v_{\mathcal{J}^c}^n | \mathcal{C}_{(m_{\mathcal{I}}, m'_{\mathcal{I}^c})} = c \right) 1 \left\{ (v_{\mathcal{J}}^n(m_{\mathcal{J}}), v_{\mathcal{J}^c}^n, y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \right\} \end{aligned} \quad (150)$$

$$= \sum_{v_{\mathcal{J}^c}^n} \prod_{j \in \mathcal{J}^c} \prod_{i=1}^n p_{V_j | V_{\mathcal{A}_j}} \left(v'_{j,i} | v_{\mathcal{A}_j \cap \mathcal{J}, i}(m_{\mathcal{J}}), v'_{\mathcal{A}_j \cap \mathcal{J}^c, i} \right) 1 \left\{ (v_{\mathcal{J}}^n(m_{\mathcal{J}}), v_{\mathcal{J}^c}^n, y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \right\} \quad (151)$$

$$\leq 2^{n(H(V_{\mathcal{J}^c} | Y_k V_{\mathcal{J}}) - \sum_{j \in \mathcal{J}^c} H(V_j | V_{\mathcal{A}_j}) + (|\mathcal{J}^c| + 1)\delta(\epsilon'))}, \quad (152)$$

where $\delta(\epsilon')$ is a term that tends to zero as $\epsilon' \rightarrow 0$, and (152) follows from the fact that $\prod_{i=1}^n p_{V_j | V_{\mathcal{A}_j}}(v_{j,i} | v_{\mathcal{A}_j, i}) \leq 2^{-n(H(V_j | V_{\mathcal{A}_j}) - \delta(\epsilon'))}$ for any $(v_j^n, v_{\mathcal{A}_j}^n) \in \mathcal{T}_{\epsilon'}^{(n)}$ and $\left| \left\{ v_{\mathcal{J}^c}^n : (v_{\mathcal{J}}^n, v_{\mathcal{J}^c}^n, y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \right\} \right| \leq 2^{n(H(V_{\mathcal{J}^c} | Y_k V_{\mathcal{J}}) + \delta(\epsilon'))}$ for any $(y_k^n, v_{\mathcal{J}}^n)$.

Combining (143), (148) and (152) gives

$$\begin{aligned} & \mathbb{P} \left((V_{\mathcal{D}_k}^n(M_{\mathcal{I}}, m'_{\mathcal{I}^c}), Y_k^n) \in \mathcal{T}_{\epsilon'}^{(n)} \text{ for some } m'_{\mathcal{I}^c}, m'_{\mathcal{I}^c} \not\Leftarrow M_{\mathcal{I}^c} \right) \\ & \leq 2^{n(\sum_{j \in \mathcal{J}^c} r_j - (\sum_{j \in \mathcal{J}^c} H(V_j | V_{\mathcal{A}_j}) - H(V_{\mathcal{J}^c} | Y_k V_{\mathcal{J}}) - (|\mathcal{J}^c| + 1)\delta(\epsilon')))}. \end{aligned} \quad (153)$$

Hence if $\sum_{j \in \mathcal{J}^c} r_j < \sum_{j \in \mathcal{J}^c} H(V_j | V_{\mathcal{A}_j}) - H(V_{\mathcal{J}^c} | Y_k V_{\mathcal{J}}) - (|\mathcal{J}^c| + 1)\delta(\epsilon')$ for all $\mathcal{J} \in \mathbb{J}$, then the third term of (133) tends to zero as $n \rightarrow \infty$. Letting ϵ' small enough, this completes the proof of the inner bound.

Besides, it is worth noting that although Multivariate Packing Lemma (Lemma 2) has not been employed directly in the proof, the derivation after (151) is essentially the same as that of Multivariate Packing Lemma.

B. Outer Bound $\mathcal{R}_1^{(o)}$

For fixed $p_{U_{[1:L]}|S}$, we first introduce a set of auxiliary random variables $U_{[1:L]}^n$ that follow distribution $\prod_{i=1}^n p_{U_{[1:L]}|S}(u_{[1:L],i}|s_i)$. Hence the Markov chains $U_{[1:L]}^n \rightarrow S^n \rightarrow X^n \rightarrow Y_k^n \rightarrow \hat{S}_k^n, 1 \leq k \leq K$ hold. Consider that

$$\begin{aligned} I(Y_{\mathcal{A}}^n; U_{\mathcal{B}}^n | U_{\mathcal{C}}^n) \\ = \sum_{t=1}^n I(Y_{\mathcal{A}}^n; U_{\mathcal{B},t} | U_{\mathcal{C}}^n U_{\mathcal{B}}^{t-1}) \end{aligned} \quad (154)$$

$$= \sum_{t=1}^n H(U_{\mathcal{B},t} | U_{\mathcal{C}}^n U_{\mathcal{B}}^{t-1}) - H(U_{\mathcal{B},t} | U_{\mathcal{C}}^n U_{\mathcal{B}}^{t-1} Y_{\mathcal{A}}^n) \quad (155)$$

$$= \sum_{t=1}^n H(U_{\mathcal{B},t} | U_{\mathcal{C},t}) - H(U_{\mathcal{B},t} | U_{\mathcal{C}}^n U_{\mathcal{B}}^{t-1} Y_{\mathcal{A}}^n) \quad (156)$$

$$= \sum_{t=1}^n I(U_{\mathcal{B},t}; U_{\mathcal{C}}^n U_{\mathcal{B}}^{t-1} Y_{\mathcal{A}}^n | U_{\mathcal{C},t}) \quad (157)$$

$$\geq \sum_{t=1}^n I(U_{\mathcal{B},t}; \hat{S}_{\mathcal{A},t} | U_{\mathcal{C},t}) \quad (158)$$

$$= nI(U_{\mathcal{B},Q}; \hat{S}_{\mathcal{A},Q} | U_{\mathcal{C},Q}) \quad (159)$$

$$= nI(U_{\mathcal{B},Q}; \hat{S}_{\mathcal{A},Q} Q | U_{\mathcal{C},Q}) \quad (160)$$

$$\geq nI(U_{\mathcal{B},Q}; \hat{S}_{\mathcal{A},Q} | U_{\mathcal{C},Q}) \quad (161)$$

$$= nI(U_{\mathcal{B}}; \hat{S}_{\mathcal{A}} | U_{\mathcal{C}}), \quad (162)$$

where the time-sharing random variable Q is defined to be uniformly distributed $[1 : n]$ and independent of all other random variables, and in (162), $U_l \triangleq U_{l,Q}, \hat{S}_k \triangleq \hat{S}_{k,Q}, 1 \leq l \leq L, 1 \leq k \leq K$.

Set $\tilde{U}_l \triangleq U_l^n, 1 \leq l \leq L$, then (162) implies the outer bound $\mathcal{R}_1^{(o)}$ holds.

C. Outer Bound $\mathcal{R}_1'^{(o)}$

To prove the outer bound $\mathcal{R}_1^{(o)}$, we only need to show $\mathcal{R}_1^{(o)} \subseteq \mathcal{R}_1'^{(o)}$. Substituting $\mathcal{A} = \mathcal{A}_i, \mathcal{B} = \mathcal{B}_i, \mathcal{C} = \cup_{j=0}^{i-1} \mathcal{B}_j$ into $\mathcal{R}_1^{(o)}$ gives us $I(\hat{S}_{\mathcal{A}_i}; U_{\mathcal{B}_i} | U_{\cup_{j=0}^{i-1} \mathcal{B}_j}) \leq \frac{1}{n} I(Y_{\mathcal{A}_i}^n; \tilde{U}_{\mathcal{B}_i} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j})$. Summate both sides of this inequality through $i = 1$ to m , then we get

$$\frac{1}{n} \sum_{i=1}^m I(Y_{\mathcal{A}_i}^n; \tilde{U}_{\mathcal{B}_i} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j}) \geq \sum_{i=1}^m I(\hat{S}_{\mathcal{A}_i}; U_{\mathcal{B}_i} | U_{\cup_{j=0}^{i-1} \mathcal{B}_j}). \quad (163)$$

Now, we turn to upper-bounding $\frac{1}{n} \sum_{i=1}^m I(Y_{\mathcal{A}_i}^n; \tilde{U}_{\mathcal{B}_i} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j})$. We will show for any $m \geq 1$,

$$\sum_{i=1}^m I(Y_{\mathcal{A}_i}^n; \tilde{U}_{\mathcal{B}_i} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j}) \leq \sum_{i=1}^m \sum_{t=1}^n I(Y_{\mathcal{A}_i,t}; \tilde{U}_{\mathcal{B}_i} \tilde{Y}_{\mathcal{A}_{i+1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_i}^{t-1} \tilde{Y}_{\mathcal{A}_{i-1}}^{t-1}) - \sum_{t=1}^n I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m+1}}^{t-1} | \tilde{U}_{\cup_{j=0}^m \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) \quad (164)$$

by induction method, where

$$\tilde{Y}_{\mathcal{A}_i}^{t-1} \triangleq \begin{cases} Y_{\mathcal{A}_i}^{t-1}, & \text{if } i \text{ is odd;} \\ Y_{\mathcal{A}_i,t+1}^n, & \text{if } i \text{ is even.} \end{cases} \quad (165)$$

For $m = 1$,

$$\begin{aligned} & I(Y_{\mathcal{A}_1}^n; \tilde{U}_{\mathcal{B}_1} | \tilde{U}_{\mathcal{B}_0}) \\ &= \sum_{t=1}^n I(Y_{\mathcal{A}_1,t}; \tilde{U}_{\mathcal{B}_1} | \tilde{U}_{\mathcal{B}_0} \tilde{Y}_{\mathcal{A}_1}^{t-1}) \end{aligned} \quad (166)$$

$$= \sum_{t=1}^n I(Y_{\mathcal{A}_1,t}; \tilde{U}_{\mathcal{B}_1} \tilde{Y}_{\mathcal{A}_2}^{t-1} | \tilde{U}_{\mathcal{B}_0} \tilde{Y}_{\mathcal{A}_1}^{t-1}) - I(Y_{\mathcal{A}_1,t}; \tilde{Y}_{\mathcal{A}_2}^{t-1} | \tilde{U}_{\cup_{j=0}^1 \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_1}^{t-1}). \quad (167)$$

Hence (164) holds for $m = 1$.

Assume (164) holds for $m - 1$, then we have

$$\begin{aligned} & \sum_{i=1}^m I(Y_{\mathcal{A}_i}^n; \tilde{U}_{\mathcal{B}_i} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j}) \\ & \leq I(Y_{\mathcal{A}_m}^n; \tilde{U}_{\mathcal{B}_m} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j}) + \sum_{i=1}^{m-1} \sum_{t=1}^n I(Y_{\mathcal{A}_i,t}; \tilde{U}_{\mathcal{B}_i} \tilde{Y}_{\mathcal{A}_{i+1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_i}^{t-1} \tilde{Y}_{\mathcal{A}_{i-1}}^{t-1}) \\ & \quad - \sum_{t=1}^n I(Y_{\mathcal{A}_{m-1},t}; \tilde{Y}_{\mathcal{A}_m}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1}). \end{aligned} \quad (168)$$

Considering the first term of (168), we have

$$\begin{aligned} & I(Y_{\mathcal{A}_m}^n; \tilde{U}_{\mathcal{B}_m} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j}) \\ &= \sum_{t=1}^n I(Y_{\mathcal{A}_m,t}; \tilde{U}_{\mathcal{B}_m} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) \end{aligned} \quad (169)$$

$$= \sum_{t=1}^n I(Y_{\mathcal{A}_m,t}; \tilde{U}_{\mathcal{B}_m} \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) - I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1} | \tilde{U}_{\cup_{j=0}^m \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) \quad (170)$$

$$\begin{aligned} &= \sum_{t=1}^n I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) + I(Y_{\mathcal{A}_m,t}; \tilde{U}_{\mathcal{B}_m} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1} \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1}) \\ & \quad - I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1} | \tilde{U}_{\cup_{j=0}^m \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) \end{aligned} \quad (171)$$

$$\begin{aligned} &= \sum_{t=1}^n I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) + I(Y_{\mathcal{A}_m,t}; \tilde{U}_{\mathcal{B}_m} \tilde{Y}_{\mathcal{A}_{m+1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1} \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1}) \\ & \quad - I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m+1}}^{t-1} | \tilde{U}_{\cup_{j=0}^m \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1} \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1}) - I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1} | \tilde{U}_{\cup_{j=0}^m \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) \end{aligned} \quad (172)$$

$$\begin{aligned} &= \sum_{t=1}^n I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) + I(Y_{\mathcal{A}_m,t}; \tilde{U}_{\mathcal{B}_m} \tilde{Y}_{\mathcal{A}_{m+1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1} \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1}) \\ & \quad - I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m+1}}^{t-1} \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1} | \tilde{U}_{\cup_{j=0}^m \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) \end{aligned} \quad (173)$$

$$\begin{aligned} & \leq \sum_{t=1}^n I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) + I(Y_{\mathcal{A}_m,t}; \tilde{U}_{\mathcal{B}_m} \tilde{Y}_{\mathcal{A}_{m+1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1} \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1}) \\ & \quad - I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m+1}}^{t-1} | \tilde{U}_{\cup_{j=0}^m \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}). \end{aligned} \quad (174)$$

Combine (168) and (174), and utilize the following identity

$$\sum_{t=1}^n I(Y_{\mathcal{A}_m,t}; \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_m}^{t-1}) = \sum_{t=1}^n I(Y_{\mathcal{A}_{m-1},t}; \tilde{Y}_{\mathcal{A}_m}^{t-1} | \tilde{U}_{\cup_{j=0}^{m-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_{m-1}}^{t-1}), \quad (175)$$

which follows from Csiszár sum identity [16, p. 25], then we have (164) holds for m . Hence (164) holds for any $m \geq 1$.

From (164), we have

$$\sum_{i=1}^m I\left(Y_{\mathcal{A}_i}^n; \tilde{U}_{\mathcal{B}_i} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j}\right) \leq \sum_{i=1}^m \sum_{t=1}^n I\left(Y_{\mathcal{A}_i, t}; \tilde{U}_{\mathcal{B}_i} \tilde{Y}_{\mathcal{A}_{i+1}}^{t-1} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_i}^{t-1} \tilde{Y}_{\mathcal{A}_{i-1}}^{t-1}\right) \quad (176)$$

$$= n \sum_{i=1}^m I\left(Y_{\mathcal{A}_i, Q}; \tilde{U}_{\mathcal{B}_i} \tilde{Y}_{\mathcal{A}_{i+1}}^{Q-1} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j} \tilde{Y}_{\mathcal{A}_i}^{Q-1} \tilde{Y}_{\mathcal{A}_{i-1}}^{Q-1} Q\right) \quad (177)$$

$$= n \sum_{i=1}^m I\left(Y_{\mathcal{A}_i}; \tilde{U}_{\mathcal{B}_i} \tilde{W}_{\mathcal{A}_{i+1}} | \tilde{U}_{\cup_{j=0}^{i-1} \mathcal{B}_j} \tilde{W}_{\mathcal{A}_i} \tilde{W}_{\mathcal{A}_{i-1}}\right), \quad (178)$$

where the time-sharing random variable Q is defined above, and $Y_k \triangleq Y_{k, Q}$, $W_k \triangleq (Y_k^{Q-1}, Q)$, $W'_k \triangleq (Y_{k, Q+1}^n, Q)$, $1 \leq k \leq K$, and $\tilde{W}_{\mathcal{A}_i} \triangleq W_{\mathcal{A}_i}$, if i is odd; $W'_{\mathcal{A}_i}$, otherwise.

If we redefine (165) as

$$\tilde{Y}_{\mathcal{A}_i}^{t-1} \triangleq \begin{cases} Y_{\mathcal{A}_i, t+1}^n, & \text{if } i \text{ is odd;} \\ Y_{\mathcal{A}_i}^{t-1}, & \text{if } i \text{ is even,} \end{cases} \quad (179)$$

then (178) still holds for $\tilde{W}_{\mathcal{A}_i} \triangleq W'_{\mathcal{A}_i}$, if i is odd; $W_{\mathcal{A}_i}$, otherwise.

Combine bounds (163) and (178), then the outer bound $\mathcal{R}_1^{(o)}$ holds.

D. Outer Bound $\mathcal{R}_2^{(o)}$

For fixed $p_{U_{[1:L]}|Y_{[1:K]}}$, we first introduce a set of auxiliary random variables $U_{[1:L]}^n$ that follow distribution $\prod_{i=1}^n p_{U_{[1:L]}|Y_{[1:K]}}(u_{[1:L], i} | y_{[1:K], i})$. Hence the Markov chains $S^n \rightarrow X^n \rightarrow Y_{[1:K]}^n \rightarrow U_{[1:L]}^n$ hold. Note that different from the proof of $\mathcal{R}_1^{(o)}$, the auxiliary random variables $U_{[1:L]}^n$ here is introduced at receiver sides, and

$p_{Y_{[1:K]}|U_{[1:L]}|X} = p_{U_{[1:L]}|Y_{[1:K]}}p_{Y_{[1:K]}|X}$ forms a new memoryless broadcast channel. Consider that

$$\begin{aligned} & I(S^n; Y_B^n U_{B'}^n | Y_C^n U_{C'}^n) \\ & \leq I(X^n; Y_B^n U_{B'}^n | Y_C^n U_{C'}^n) \end{aligned} \quad (180)$$

$$= \sum_{t=1}^n I(X^n; Y_{B,t} U_{B',t} | Y_C^n U_{C'}^n Y_B^{t-1} U_{B'}^{t-1}) \quad (181)$$

$$= \sum_{t=1}^n H(Y_{B,t} U_{B',t} | Y_C^n U_{C'}^n Y_B^{t-1} U_{B'}^{t-1}) - H(Y_{B,t} U_{B',t} | Y_C^n U_{C'}^n Y_B^{t-1} U_{B'}^{t-1} X^n) \quad (182)$$

$$\leq \sum_{t=1}^n H(Y_{B,t} U_{B',t} | Y_{C,t} U_{C',t}) - H(Y_{B,t} U_{B',t} | Y_{C,t} U_{C',t} X_t) \quad (183)$$

$$= \sum_{t=1}^n I(Y_{B,t} U_{B',t}; X_t | Y_{C,t} U_{C',t}) \quad (184)$$

$$= nI(Y_{B,Q} U_{B',Q}; X_Q | Y_{C,Q} U_{C',Q} Q) \quad (185)$$

$$= nH(Y_{B,Q} U_{B',Q} | Y_{C,Q} U_{C',Q} Q) - nH(Y_{B,Q} U_{B',Q} | Y_{C,Q} U_{C',Q} X_Q Q) \quad (186)$$

$$\leq nH(Y_{B,Q} U_{B',Q} | Y_{C,Q} U_{C',Q}) - nH(Y_{B,Q} U_{B',Q} | Y_{C,Q} U_{C',Q} X_Q) \quad (187)$$

$$= nI(Y_{B,Q} U_{B',Q}; X_Q | Y_{C,Q} U_{C',Q}) \quad (188)$$

$$= nI(X; Y_B U_{B'} | Y_C U_{C'}), \quad (189)$$

where (183) follows from Markov chain $U_{[1:L]}^{t-1} U_{[1:L],t+1}^n Y_{[1:K]}^{t-1} Y_{[1:K],t+1}^n X^{t-1} X_{t+1}^n \rightarrow X_t \rightarrow Y_{[1:K],t} U_{[1:L],t}$ and the fact conditioning reduces entropy, (187) follows from Markov chain $Q \rightarrow X_Q \rightarrow Y_{[1:K],Q} U_{[1:L],Q}$ and the fact conditioning reduces entropy, the time-sharing random variable Q is defined to be uniformly distributed $[1:n]$ and independent of all other random variables, and in (189), $U_l \triangleq U_{l,Q}$, $Y_k \triangleq Y_{k,Q}$, $X \triangleq X_Q$, $1 \leq l \leq L$, $1 \leq k \leq K$.

On the other hand,

$$\begin{aligned} & I(S^n; Y_B^n U_{B'}^n | Y_C^n U_{C'}^n) \\ & = \sum_{t=1}^n I(S_t; Y_B^n U_{B'}^n | Y_C^n U_{C'}^n S^{t-1}) \end{aligned} \quad (190)$$

$$= nI(S_Q; Y_B^n U_{B'}^n | Y_C^n U_{C'}^n S^{Q-1} Q) \quad (191)$$

$$= nI(S; \tilde{Y}_B \tilde{U}_{B'} | \tilde{Y}_C \tilde{U}_{C'}), \quad (192)$$

Set $S \triangleq S_Q$, $\tilde{U}_l \triangleq U_l^n S^{Q-1} Q$, $\tilde{Y}_k \triangleq Y_k^n S^{Q-1} Q$, $1 \leq l \leq L$, $1 \leq k \leq K$, then combining (189) and (192) gives us the outer bound $\mathcal{R}_2^{(o)}$.

APPENDIX D

PROOF OF THEOREM 3

A. Inner Bound

For the inner bound $\mathcal{R}^{(i)}$ of Theorem 1, retain all the random variables V_i 's corresponding to the sets $\mathcal{G}_i = [1 : K], [2 : K], \dots, \{K\}$, rename them and corresponding rates r_i 's to V'_1, V'_2, \dots, V'_K and r'_1, r'_2, \dots, r'_K , respectively, and set all the other random variables to empty, then $\mathcal{R}^{(i)}$ reduces to

$$\begin{aligned} \mathcal{R}^{(i)} = & \left\{ D_{[1:K]} : \text{There exist some pmf } p_{V'_{[1:K]}|S}, \text{ vector } r'_{[1:K]}, \right. \\ & \text{and functions } x(v'_{[1:K]}, s), \hat{s}_k(v'_{[1:k]}, y_k), 1 \leq k \leq K \text{ such that} \\ & \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\ & \sum_{j=1}^k r'_j > I(V'_{[1:k]}; S), 1 \leq k \leq K, \\ & \left. r'_k < I(V'_k; Y_k | V'_{[1:k-1]}), 1 \leq k \leq K \right\}, \end{aligned} \quad (193)$$

and the coding scheme in the proof of Theorem 1 reduces to a superposition coding scheme. Define a set of random variables $V_k \triangleq V'_{[1:k]}, 1 \leq k \leq K$. Substitute these into (193), then the inner bound of Theorem 3 is recovered.

B. Outer Bound

For the outer bound, we provide two proofs. The first follows from Theorem 1, and the second is a more simple and direct proof that does not utilize the Csiszár sum identity.

Proof method 1: Set $L = K, \mathcal{A} = \mathcal{B} = [1 : k], \mathcal{C} = [1 : k-1]$ for $1 \leq k \leq m$, and $U_0 \triangleq \emptyset, U_K \triangleq S, p_{U_{[1:K-1]}|S} = p_{U_{K-1}|S} p_{U_{K-2}|U_{K-1}} \dots p_{U_1|U_2}$. Substitute these into the outer bound $\mathcal{R}_1^{(o)}$ of Theorem 1, and utilize the degradation of the channel, then we get $I(\hat{S}'_k; U_k | U_{k-1}) \leq \frac{1}{n} I(Y^n_{[1:k]}; \tilde{U}_{[1:k]} | \tilde{U}_{[1:k-1]})$, $\hat{S}'_k \triangleq \hat{S}_{[1:k]}, 1 \leq k \leq K$. Hence $I(\hat{S}'_k; U_k | U_{k-1}) \in \frac{1}{n} \mathcal{B}_{DBC}(p_X^n p_{Y^n_{[1:K]}|X^n}) \subseteq \mathcal{B}_{DBC}(p_X p_{Y_{[1:K]}|X})$ for some $p_{\hat{S}_{[1:K]}|S}$ and p_X . The last equality follows from that for the memoryless degraded broadcast channel, the capacity region will not expand after extending the channel to n lettered one.

$$\begin{aligned} \mathcal{R}_2^{(o)} = & \left\{ D_{[1:K]} : \text{There exists some pmf } p_X \text{ and some functions } \hat{s}_k^n(\tilde{y}_k), 1 \leq k \leq K \text{ such that} \right. \\ & \mathbb{E} d_k(S, \hat{S}_k) \leq D_k, 1 \leq k \leq K, \\ & \text{and for any pmf } p_{U_{[1:L]}|Y_{[1:K]}}, \text{ one can find } p_{\tilde{Y}_{[1:K]}|S^n} p_{\tilde{U}_{[1:L]}|\tilde{Y}_{[1:K]}} \text{ satisfying} \\ & \left. \frac{1}{n} I(S^n; \tilde{Y}_{\mathcal{B}} \tilde{U}_{\mathcal{B}'} | \tilde{Y}_{\mathcal{C}} \tilde{U}_{\mathcal{C}'}) \leq I(X; Y_{\mathcal{B}} U_{\mathcal{B}'} | Y_{\mathcal{C}} U_{\mathcal{C}'}) \text{ for any } \mathcal{B}, \mathcal{C} \subseteq [1 : K], \mathcal{B}', \mathcal{C}' \subseteq [1 : L] \right\}. \end{aligned} \quad (194)$$

In addition, set $\mathcal{B} = [1 : k], \mathcal{C} = [1 : k-1]$ for $1 \leq k \leq K$ and $\mathcal{B}' = \mathcal{C}' = \emptyset$. Substitute these into the outer bound $\mathcal{R}_2^{(o)}$ of Theorem 1, and utilize the degradation of the channel, then we get

$$\frac{1}{n} I(S^n; \tilde{Y}_{[1:k]} | \tilde{Y}_{[1:k-1]}) \leq I(X; Y_k | Y_{k-1}). \quad (195)$$

Summate both sides of this inequality through $k = 1$ to K , then we get

$$\frac{1}{n} I(S^n; \tilde{Y}_{[1:k]}) \leq I(X; Y_k). \quad (196)$$

In addition,

$$\frac{1}{n} I(S^n; \tilde{Y}_{[1:k]}) \geq \frac{1}{n} I(S^n; \hat{S}_{[1:k]}^n) \quad (197)$$

$$= \frac{1}{n} \sum_{t=1}^n I(S_t; \hat{S}_{[1:k]}^n | S^{t-1}) \quad (198)$$

$$= \frac{1}{n} \sum_{t=1}^n I(S_t; \hat{S}_{[1:k]}^n S^{t-1}) \quad (199)$$

$$\geq \frac{1}{n} \sum_{t=1}^n I(S_t; \hat{S}_{[1:k],t}) \quad (200)$$

$$= I(S_Q; \hat{S}_{[1:k],Q} | Q) \quad (201)$$

$$= I(S_Q; \hat{S}_{[1:k],Q} Q) \quad (202)$$

$$\geq I(S_Q; \hat{S}_{[1:k],Q}) \quad (203)$$

$$= I(S; \hat{S}_{[1:k]}) \quad (204)$$

$$= I(S; \hat{S}'_k), \quad (205)$$

where the time-sharing random variable Q is defined to be uniformly distributed $[1 : n]$ and independent of all other random variables, and $S \triangleq S_Q, \hat{S}_k \triangleq \hat{S}_{k,Q}, \hat{S}'_k \triangleq \hat{S}_{[1:k]}, 1 \leq k \leq K$. Combining (196) and (205) gives us $I(S; \hat{S}'_k) \leq I(X; Y_k)$. Hence the outer bound $\mathcal{R}^{(o)}$ of Theorem 3 holds.

Proof method 2: For fixed $p_{U_{K-1}|S} p_{U_{K-2}|U_{K-1}} \cdots p_{U_1|U_2}$, we first introduce a set of auxiliary random variables $U_{[1:K-1]}^n$ that follow distribution $\prod_{i=1}^n p_{U_{K-1}|S}(u_{K-1,i}|s_i) p_{U_{K-2}|U_{K-1}}(u_{K-2,i}|u_{K-1,i}) \cdots p_{U_1|U_2}(u_{1,i}|u_{2,i})$. Then $U_1^n \rightarrow U_2^n \rightarrow \cdots \rightarrow U_{K-1}^n \rightarrow S^n \rightarrow X^n \rightarrow Y_K^n \rightarrow Y_{K-1}^n \rightarrow \cdots \rightarrow Y_1^n$ follows a Markov chain. We first derive a lower bound for $I(Y_k^n; U_k^n | U_{k-1}^n)$.

$$\begin{aligned}
& I(Y_k^n; U_k^n | U_{k-1}^n) \\
&= I(Y_{[1:k]}^n; U_k^n | U_{k-1}^n)
\end{aligned} \tag{206}$$

$$= \sum_{i=1}^n I(Y_{[1:k]}^n; U_{k,i} | U_{k+1}^n U_k^{i-1}) \tag{207}$$

$$= \sum_{i=1}^n H(U_{k,i} | U_{k-1}^n U_k^{i-1}) - H(U_{k,i} | U_{k-1}^n U_k^{i-1} Y_{[1:k]}^n) \tag{208}$$

$$= \sum_{i=1}^n H(U_{k,i} | U_{k-1,i}) - H(U_{k,i} | U_{k-1}^n U_k^{i-1} Y_{[1:k]}^n) \tag{209}$$

$$= \sum_{i=1}^n I(U_{k,i}; U_{k-1}^n U_k^{i-1} Y_{[1:k]}^n | U_{k-1,i}) \tag{210}$$

$$\geq \sum_{i=1}^n I(U_{k,i}; \hat{S}_{[1:k],i} | U_{k-1,i}) \tag{211}$$

$$= nI(U_{k,Q}; \hat{S}_{[1:k],Q} | U_{k-1,Q} Q) \tag{212}$$

$$= nI(U_{k,Q}; \hat{S}_{[1:k],Q} Q | U_{k-1,Q}) \tag{213}$$

$$\geq nI(U_{k,Q}; \hat{S}_{[1:k],Q} | U_{k-1,Q}) \tag{214}$$

$$= nI(U_k; \hat{S}_{[1:k]} | U_{k-1}), \tag{215}$$

where the time-sharing random variable Q is defined to be uniformly distributed $[1 : n]$ and independent of all other random variables, and in (215), $\hat{S}_k \triangleq \hat{S}_{k,Q}$, $U_k \triangleq U_{k,Q}$, $1 \leq k \leq K$.

Next, we turn to upper-bounding $I(Y_k^n; U_k^n | U_{k-1}^n)$, and write the following:

$$\begin{aligned}
& I(Y_k^n; U_k^n | U_{k-1}^n) \\
&= \sum_{i=1}^n I(Y_{k,i}; U_k^n | U_{k-1}^n Y_k^{i-1})
\end{aligned} \tag{216}$$

$$\leq \sum_{i=1}^n I(Y_{k,i}; U_k^n Y_{k+1}^{i-1} | U_{k-1}^n Y_k^{i-1}) \tag{217}$$

$$= n \sum_{i=1}^n I(Y_{k,Q}; U_k^n Y_{k+1}^{Q-1} | U_{k-1}^n Y_k^{Q-1} Q) \tag{218}$$

$$= nI(Y_k; V'_k | V'_{k-1}), \tag{219}$$

where the time-sharing random variable Q is defined above, and $Y_{K+1}^{Q-1} \triangleq X^{Q-1}$, $V'_k \triangleq (U_k^n, Y_{k+1}^{Q-1}, Q)$, $Y_k \triangleq Y_{k,Q}$, $1 \leq k \leq K$.

Obviously, $V'_1 \rightarrow V'_2 \rightarrow \dots \rightarrow V'_K \rightarrow X \rightarrow Y_k$ forms a Markov chain, hence

$$(I(Y_k; V'_k | V'_{k-1}) : k \in [1 : K]) \in \mathcal{B}_{DBC}(p_X p_{Y_{[1:K]} | X}). \tag{220}$$

Combining this with (215) and (219), we have

$$\left(I \left(U_k; \hat{S}_{[1:k]} | U_{k-1} \right) : k \in [1 : K] \right) \in \mathcal{B}_{DBC} \left(p_X p_{Y_{[1:K]} | X} \right). \quad (221)$$

On the other hand, it can be proved that $nI \left(\hat{S}_{[1:k]}; S \right) \leq I \left(\hat{S}_{[1:k]}^n; S^n \right) \leq I \left(Y_k^n; X^n \right) \leq nI \left(Y_k; X \right)$, i.e., $I \left(S; \hat{S}_k' \right) \leq I \left(X; Y_k \right)$. Combining it with (221) completes the proof.

APPENDIX E

PROOF OF THEOREM 8

For Wyner-Ziv Gaussian broadcast with bandwidth mismatch case (bandwidth mismatch factor b), Theorem 7 states that if $D_{[1:K]}$ is achievable, then there exists some pmf $p_{V_K | S} p_{V_{K-1} | V_K} \cdots p_{V_1 | V_2}$ and functions $\hat{s}_k(v_k, z_k)$, $1 \leq k \leq K$ such that

$$\mathbb{E} d \left(S, \hat{S}_k \right) \leq D_k, \quad (222)$$

and for any pmf $p_{U_{K-1} | S} p_{U_{K-2} | U_{K-1}} \cdots p_{U_1 | U_2}$,

$$\frac{1}{b} \left(I \left(V_k; U_k | U_{k-1} Z_k \right) : k \in [1 : K] \right) \in \mathcal{C}_{GBC} \quad (223)$$

holds, where the capacity of Gaussian broadcast channel \mathcal{C}_{GBC} is given in (57).

Choose $U_{K-1} = S + E'_{K-1}$ and $U_k = U_{k+1} + E'_k$, $1 \leq k \leq K-2$, where $E'_k \sim \mathcal{N}(0, \tau'_k)$ is independent of all the other random variables. Define $E_k = \sum_{j=k}^{K-1} E'_j \sim \mathcal{N}(0, \tau_k)$ with $\tau_k = \sum_{j=k}^{K-1} \tau'_j$. Then

$$I(V_1; U_1 | Z_1) \geq I(\hat{S}_1; U_1 | Z_1) \quad (224)$$

$$= h(U_1 | Z_1) - h(U_1 | \hat{S}_1 Z_1) \quad (225)$$

$$= h(U_1 | Z_1) - h(U_1 - \hat{S}_1 | \hat{S}_1 Z_1) \quad (226)$$

$$\geq h(U_1 | Z_1) - h(U_1 - \hat{S}_1) \quad (227)$$

$$\geq \frac{1}{2} \log(2\pi e(\beta_1 + \tau_1)) - \frac{1}{2} \log(2\pi e(D_1 + \tau_1)) \quad (228)$$

$$= \frac{1}{2} \log \frac{\beta_1 + \tau_1}{D_1 + \tau_1}, \quad (229)$$

where (228) follows from Gaussian distribution maximizes the differential entropy for a given second moment.

On the other hand,

$$I(V_k; U_k | U_{k-1} Z_k) \geq I(\hat{S}_k; U_k | U_{k-1} Z_k) \quad (230)$$

$$= I(\hat{S}_k; U_k | Z_k) - I(\hat{S}_k; U_{k-1} | Z_k) \quad (231)$$

$$= h(U_k | Z_k) - h(U_{k-1} | Z_k) + h(U_{k-1} | Z_k \hat{S}_k) - h(U_k | Z_k \hat{S}_k). \quad (232)$$

The first two terms of (232)

$$h(U_k | Z_k) - h(U_{k-1} | Z_k) = \frac{1}{2} \log \frac{\beta_k + \tau_k}{\beta_k + \tau_{k-1}}. \quad (233)$$

The last two terms of (232)

$$h(U_{k-1}|Z_k\hat{S}_k) - h(U_k|Z_k\hat{S}_k) = h(U_{k-1}|Z_k\hat{S}_k) - h(U_k|Z_k\hat{S}_kE'_{k-1}) \quad (234)$$

$$= h(U_{k-1}|Z_k\hat{S}_k) - h(U_{k-1}|Z_k\hat{S}_kE'_{k-1}) \quad (235)$$

$$= I(U_{k-1}; E'_{k-1}|Z_k\hat{S}_k) \quad (236)$$

$$= h(E'_{k-1}) - h(E'_{k-1}|Z_k\hat{S}_kU_{k-1}) \quad (237)$$

$$= h(E'_{k-1}) - h(E'_{k-1}|Z_k, \hat{S}_k, U_{k-1} - \hat{S}_k) \quad (238)$$

$$\geq h(E'_{k-1}) - h(E'_{k-1}|U_{k-1} - \hat{S}_k) \quad (239)$$

$$= I(E'_{k-1}; S - \hat{S}_k + E_k + E'_{k-1}) \quad (240)$$

$$\geq \frac{1}{2} \log \frac{D_k + \tau_{k-1}}{D_k + \tau_k}, \quad (241)$$

where (241) is by applying the mutual information game result that Gaussian noise is the worst additive noise under a variance constraint [14, p. 298, Problem 9.21] and taking E'_{k-1} as channel input.

Combining (232), (233) and (241), we have

$$I(V_k; U_k|U_{k-1}Z_k) \geq \frac{1}{2} \log \frac{(\beta_k + \tau_k)(D_k + \tau_{k-1})}{(\beta_k + \tau_{k-1})(D_k + \tau_k)}. \quad (242)$$

(223), (229) and (242) imply Theorem 8 holds.

APPENDIX F

PROOF OF THEOREM 9

Observe that if there is no information transmitted over the channel, receiver k could produce a reconstruction within distortion β_k . Hence we only need consider the case of $D_{[1:K]}$ with

$$D_k \leq \beta_k, 1 \leq k \leq K. \quad (243)$$

For Wyner-Ziv binary broadcast with bandwidth mismatch case (bandwidth mismatch factor b), Theorem 7 states that if $D_{[1:K]}$ is achievable, then there exists some pmf $p_{V_K|S}p_{V_{K-1}|V_K} \cdots p_{V_1|V_2}$ and functions $\hat{s}_k(v_k, z_k), 1 \leq k \leq K$ such that

$$\mathbb{E}d(S, \hat{S}_k) = \mathbb{P}(\hat{S}_k \oplus S = 1) \leq D_k, \quad (244)$$

and for any pmf $p_{U_{K-1}|S}p_{U_{K-2}|U_{K-1}} \cdots p_{U_1|U_2}$,

$$\frac{1}{b} (I(V_k; U_k|U_{k-1}Z_k) : k \in [1:K]) \in \mathcal{C}_{BBC} \quad (245)$$

holds, where the capacity of binary broadcast channel \mathcal{C}_{BBC} is given in (62) [13].

Define the sets

$$\mathcal{A}_k = \{v_k : \hat{s}_k(v_k, 0) = \hat{s}_k(v_k, 1)\}, 1 \leq k \leq K, \quad (246)$$

so that their complements

$$\mathcal{A}_k^c = \{v_k : \hat{s}_k(v_k, 0) \neq \hat{s}_k(v_k, 1)\}, 1 \leq k \leq K. \quad (247)$$

By hypothesis,

$$\mathbb{E}d(S, \hat{S}_k) = \mathbb{P}(V_k \in \mathcal{A}_k) \mathbb{E}[d(S, \hat{S}_k) | V_k \in \mathcal{A}_k] + \mathbb{P}(V_k \in \mathcal{A}_k^c) \mathbb{E}[d(S, \hat{S}_k) | V_k \in \mathcal{A}_k^c] \quad (248)$$

$$\leq D_k. \quad (249)$$

We first show that

$$\mathbb{E}[d(S, \hat{S}_k) | V_k \in \mathcal{A}_k^c] \geq \beta_k. \quad (250)$$

To do this, we write

$$\mathbb{E}[d(S, \hat{S}_k) | V_k \in \mathcal{A}_k^c] = \sum_{v_k \in \mathcal{A}_k^c} \frac{\mathbb{P}(V_k = v_k)}{\mathbb{P}(V_k \in \mathcal{A}_k^c)} \mathbb{E}[d(S, \hat{S}_k) | V_k = v_k]. \quad (251)$$

If $v_k \in \mathcal{A}_k^c$ and $\hat{s}_k(v_k, 0) = 0$ then $\hat{s}_k(v_k, 1) = 1$. Therefore, for such v_k ,

$$\mathbb{E}[d(S, \hat{S}_k) | V_k = v_k] = \mathbb{P}(Z_k = 0, S = 1 | V_k = v_k) + \mathbb{P}(Z_k = 1, S = 0 | V_k = v_k) \quad (252)$$

$$= \mathbb{P}(S = 1 | V_k = v_k) \mathbb{P}(Z_k = 0 | S = 1) + \mathbb{P}(S = 0 | V_k = v_k) \mathbb{P}(Z_k = 1 | S = 0) \quad (253)$$

$$= \beta_k, \quad (254)$$

where (253) follows that $Z_k \rightarrow S \rightarrow V_k$ forms a Markov chain. If $v_k \in \mathcal{A}_k^c$ but $\hat{s}_k(v_k, 0) = 1$, then for such v_k ,

$$\mathbb{E}[d(S, \hat{S}_k) | V_k = v_k] = 1 - \beta_k \geq \beta_k, \quad (255)$$

since $\beta_k \leq \frac{1}{2}$. Therefore, (250) follows from (254) and (255).

Now we write

$$\mathbb{E}[d(S, \hat{S}_k) | V_k \in \mathcal{A}_k] = \sum_{v_k \in \mathcal{A}_k} \frac{\mathbb{P}(V_k = v_k)}{\mathbb{P}(V_k \in \mathcal{A}_k)} \mathbb{E}[d(S, \hat{S}_k) | V_k = v_k], \quad (256)$$

and define $g_k(v_k) \triangleq \hat{s}_k(v_k, 0)$, $\lambda_{v_k} \triangleq \frac{\mathbb{P}(V_k = v_k)}{\mathbb{P}(V_k \in \mathcal{A}_k)}$, $\mu_k \triangleq \mathbb{P}(V_k \in \mathcal{A}_k)$,

$$d_{v_k} \triangleq \mathbb{E}[d(S, \hat{S}_k) | V_k = v_k] = \mathbb{P}(S \neq g_k(v_k) | V_k = v_k), \quad (257)$$

then utilizing (249) and (250), we have

$$d'_k \triangleq \mu_k \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} d_{v_k} + (1 - \mu_k) \beta_k \leq D_k. \quad (258)$$

Next we will show

$$I(V_k; U_k | Z_k) \geq \frac{\beta_k - D_k}{\beta_k - \alpha_k} (H_2(\beta_k \star \tau_k) - (H_4(\alpha_k, \beta_k, \tau_k) - H_2(\alpha_k \star \beta_k))). \quad (259)$$

Choose $U_{K-1} = S \oplus E'_{K-1}$ and $U_k = U_{k+1} \oplus E'_k$, $1 \leq k \leq K-2$, where $E'_k \sim \text{Bern}(\tau'_k)$ is independent of all the other random variables. Define $E_k = E'_{K-1} \oplus E'_{K-2} \oplus \dots \oplus E'_k \sim \text{Bern}(\tau_k)$ with $\tau_k = \tau'_{K-1} \star \tau'_{K-2} \star \dots \star \tau'_k$.

Then

$$I(V_k; U_k | Z_k) = H(U_k | Z_k) - H(U_k | V_k, Z_k) \quad (260)$$

$$\begin{aligned} &= H_2(\beta_k \star \tau_k) - \mu_k \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} H(U_k | Z_k, V_k = v_k) \\ &\quad - (1 - \mu_k) \sum_{v_k \in \mathcal{A}_k^c} \frac{\mathbb{P}(V_k = v_k)}{\mathbb{P}(V_k \in \mathcal{A}_k^c)} H(U_k | Z_k, V_k = v_k). \end{aligned} \quad (261)$$

For fixed v_k , define a set of random variables $(V'_k, S', U'_k, Z'_k) \sim 1\{v'_k = v_k\} p_{SU_k Z_k | V_k}(s', u'_k, z'_k | v'_k)$, then $H(U'_k Z'_k | V'_k) = H(U_k Z_k | V_k = v_k)$ and $H(Z'_k | V'_k) = H(Z_k | V_k = v_k)$. Since $p_{SU_k Z_k | V_k}$ satisfies

$$p_{SU_k Z_k | V_k}(s', u'_k, z'_k | v'_k) = p_{S | V_k}(s' | v'_k) p_{Z_k | S}(z'_k | s') p_{U_k | S}(u'_k | s'), \quad (262)$$

it holds that $Z'_k = S' \oplus B_k, U'_k = S' \oplus E_k$. Hence $Z'_k \oplus U'_k = B_k \oplus E_k$.

For fixed v_k , consider

$$H(U_k | Z_k, V_k = v_k) = H(U_k Z_k | V_k = v_k) - H(Z_k | V_k = v_k) \quad (263)$$

$$= H(U'_k Z'_k | V'_k) - H(Z'_k | V'_k) \quad (264)$$

$$= H(U'_k | Z'_k V'_k) \quad (265)$$

$$= H(U'_k \oplus Z'_k | Z'_k V'_k) \quad (266)$$

$$= H(B_k \oplus E_k | Z'_k V'_k) \quad (267)$$

$$\leq H(B_k \oplus E_k) \quad (268)$$

$$= H_2(\beta_k \star \tau_k). \quad (269)$$

Combine (261) and (269), then it holds that

$$I(V_k; U_k | Z_k) \geq H_2(\beta_k \star \tau_k) - \mu_k \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} H(U_k | Z_k, V_k = v_k) - (1 - \mu_k) H_2(\beta_k \star \tau_k) \quad (270)$$

$$= \mu_k H_2(\beta_k \star \tau_k) - \mu_k \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} H(U_k | Z_k, V_k = v_k). \quad (271)$$

Now we consider the second term of (271).

$$\sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} H(U_k | Z_k, V_k = v_k) = \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} (H(U_k Z_k | V_k = v_k) - H(Z_k | V_k = v_k)) \quad (272)$$

$$= \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} (H_4(d_{v_k}, \beta_k, \tau_k) - H_2(d_{v_k} \star \beta_k)) \quad (273)$$

$$= \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} G_1(d_{v_k}, \beta_k, \tau_k), \quad (274)$$

where the function $H_4(x, y, z)$ is defined in (85) and

$$G_1(x, y, z) \triangleq H_4(x, y, z) - H_2(x \star y). \quad (275)$$

Equality (273) follows from calculating the entropies according to the definition.

Now we show that $G_1(x, y, z)$ is concave in x . To do this, we consider

$$\begin{aligned} & \frac{\partial^2}{\partial x^2} G_1(x, y, z) \\ &= -\frac{(yz - \bar{y}\bar{z})^2}{xyz + \bar{x}\bar{y}\bar{z}} - \frac{(\bar{y}z - y\bar{z})^2}{x\bar{y}z + \bar{x}y\bar{z}} - \frac{(y\bar{z} - \bar{y}z)^2}{xy\bar{z} + \bar{x}\bar{y}z} - \frac{(\bar{y}\bar{z} - yz)^2}{x\bar{y}\bar{z} + \bar{x}yz} + \frac{(y - \bar{y})^2}{x\bar{y} + \bar{x}y} + \frac{(y - \bar{y})^2}{xy + \bar{x}\bar{y}} \end{aligned} \quad (276)$$

$$= -\left(\frac{(yz - \bar{y}\bar{z})^2}{xyz + \bar{x}\bar{y}\bar{z}} + \frac{(\bar{y}z - y\bar{z})^2}{x\bar{y}z + \bar{x}y\bar{z}} - \frac{(y - \bar{y})^2}{xy + \bar{x}\bar{y}} \right) - \left(\frac{(\bar{y}z - y\bar{z})^2}{x\bar{y}\bar{z} + \bar{x}yz} + \frac{(\bar{y}\bar{z} - yz)^2}{x\bar{y}\bar{z} + \bar{x}yz} - \frac{(y - \bar{y})^2}{x\bar{y} + \bar{x}y} \right) \quad (277)$$

$$\leq 0, \quad (278)$$

where (278) follows from the following inequality

$$\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} = \frac{1}{b_1 + b_2} (b_1 + b_2) \left(\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} \right) \quad (279)$$

$$= \frac{1}{b_1 + b_2} \left(a_1^2 + a_2^2 + \frac{b_2 a_1^2}{b_1} + \frac{b_1 a_2^2}{b_2} \right) \quad (280)$$

$$\geq \frac{1}{b_1 + b_2} (a_1^2 + a_2^2 + 2a_1 a_2) \quad (281)$$

$$= \frac{(a_1 + a_2)^2}{b_1 + b_2}, \quad (282)$$

for $b_1, b_2 > 0$ and arbitrary real numbers a_1, a_2 . (278) implies $G_1(x, y, z)$ is concave in x .

Then combining the concavity of $G_1(x, y, z)$ with (271) and (274), we have

$$I(V_k; U_k | Z_k) \geq \mu_k \left(H_2(\beta_k \star \tau_k) - G_1 \left(\sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} d_{v_k}, \beta_k, \tau_k \right) \right) \quad (283)$$

$$= \mu_k (H_2(\beta_k \star \tau_k) - G_1(\alpha_k, \beta_k, \tau_k)) \quad (284)$$

where

$$\alpha_k \triangleq \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} d_{v_k}. \quad (285)$$

From (258), α_k satisfies

$$\mu_k \alpha_k + (1 - \mu_k) \beta_k \leq D_k. \quad (286)$$

Combine (286) with $D_k \leq \beta_k$ (i.e., (243)), then we have

$$0 \leq \alpha_k \leq D_k \leq \beta_k. \quad (287)$$

Therefore,

$$I(V_k; U_k | Z_k) \geq \frac{\beta_k - D_k}{\beta_k - \alpha_k} (H_2(\beta_k \star \tau_k) - G_1(\alpha_k, \beta_k, \tau_k)) \quad (288)$$

$$= \frac{\beta_k - D_k}{\beta_k - \alpha_k} (H_2(\beta_k \star \tau_k) - (H_4(\alpha_k, \beta_k, \tau_k) - H_2(\alpha_k \star \beta_k))), \quad (289)$$

i.e., (259) holds.

Next we will show

$$I(V_k; U_k | U_{k-1} Z_k) \geq \frac{\beta_k - D_k}{\beta_k - \alpha_k} (H_2(\beta_k \star \tau_k) - H_2(\beta_k \star \tau_{k-1}) - (H_4(\alpha_k, \beta_k, \tau_k) - H_4(\alpha_k, \beta_k, \tau_{k-1}))). \quad (290)$$

Consider

$$I(V_k; U_k | U_{k-1} Z_k) = H(U_k | U_{k-1} Z_k) - H(U_k | U_{k-1} Z_k V_k) \quad (291)$$

$$= H(U_{k-1} | U_k) + H(U_k | Z_k) - H(U_{k-1} | Z_k) - H(U_k | U_{k-1} Z_k V_k) \quad (292)$$

$$= H_2(\tau'_{k-1}) + H_2(\beta_k \star \tau_k) - H_2(\beta_k \star \tau_{k-1}) - H(U_k | U_{k-1} Z_k V_k). \quad (293)$$

Write the last term as

$$\begin{aligned} & H(U_k | U_{k-1} Z_k V_k) \\ &= -(1 - \mu_k) \sum_{v_k \in \mathcal{A}_k^c} \frac{\mathbb{P}(V_k = v_k)}{\mathbb{P}(V_k \in \mathcal{A}_k^c)} H(U_k | U_{k-1}, Z_k, V_k = v_k) - \mu_k \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} H(U_k | U_{k-1}, Z_k, V_k = v_k). \end{aligned} \quad (294)$$

For fixed v_k , define $(V'_k, S', U'_k, U'_{k-1}, Z'_k) \sim 1\{v'_k = v_k\} p_{SU_k U_{k-1} Z_k | V_k}(s', u'_k, u'_{k-1}, z'_k | v'_k)$. Since

$$p_{SU_k U_{k-1} Z_k | V_k}(s', u'_k, u'_{k-1}, z'_k | v'_k) = p_{S|V_k}(s' | v'_k) p_{Z_k|S}(z'_k | s') p_{U_k|S}(u'_k | s') p_{U_{k-1}|U_k}(u'_{k-1} | u'_k), \quad (295)$$

we have $Z'_k = S' \oplus B_k$, $U'_k = S' \oplus E_k$, $U'_{k-1} = U'_k \oplus E'_{k-1}$. Hence $Z'_k \oplus U'_k = B_k \oplus E_k$, $Z'_k \oplus U'_{k-1} = B_k \oplus E_{k-1}$.

Similar to the derivation for $H(U_k | U_{k-1}, V_k = v_k)$, we can write

$$H(U_k | U_{k-1}, Z_k, V_k = v_k) = H(U'_k | U'_{k-1} Z'_k V'_k) \quad (296)$$

$$= H(U'_k \oplus Z'_k | U'_{k-1} \oplus Z'_k, Z'_k, V'_k) \quad (297)$$

$$\leq H(U'_k \oplus Z'_k | U'_{k-1} \oplus Z'_k) \quad (298)$$

$$= H(B_k \oplus E_k | B_k \oplus E_{k-1}) \quad (299)$$

$$= H(B_k \oplus E_k) + H(B_k \oplus E_{k-1} | B_k \oplus E_k) - H(B_k \oplus E_{k-1}) \quad (300)$$

$$= H_2(\tau'_{k-1}) + H_2(\beta_k \star \tau_k) - H_2(\beta_k \star \tau_{k-1}). \quad (301)$$

Combine (293), (294) and (301), then we have

$$I(V_k; U_k | U_{k-1} Z_k) \geq \mu_k (H_2(\tau'_{k-1}) + H_2(\beta_k \star \tau_k) - H_2(\beta_k \star \tau_{k-1})) - \mu_k \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} H(U_k | U_{k-1}, Z_k, V_k = v_k). \quad (302)$$

Consider the last term of (302),

$$\begin{aligned} & \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} H(U_k | U_{k-1}, Z_k, V_k = v_k) \\ &= \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} (H(U_k | Z_k, V_k = v_k) + H(U_{k-1} | U_k, Z_k, V_k = v_k) - H(U_{k-1} | Z_k, V_k = v_k)) \end{aligned} \quad (303)$$

$$= \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} (H(U_k, Z_k | V_k = v_k) + H_2(\tau'_{k-1}) - H(U_{k-1}, Z_k | V_k = v_k)) \quad (304)$$

$$= H_2(\tau'_{k-1}) + \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} (H_4(d_{v_k}, \beta_k, \tau_k) - H_4(d_{v_k}, \beta_k, \tau_{k-1})) \quad (305)$$

$$= H_2(\tau'_{k-1}) + \sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} G_2(d_{v_k}, \beta_k, \tau_k, \tau_{k-1}), \quad (306)$$

where (305) is by directly calculating the entropies according to the definition, and

$$G_2(x, y, z, t) \triangleq H_4(x, y, z) - H_4(x, y, t). \quad (307)$$

Note that function $G_1(x, y, z)$ is a special case of function $G_2(x, y, z, t)$ given $t = \frac{1}{2}$, i.e.,

$$G_1(x, y, z) = G_2\left(x, y, z, \frac{1}{2}\right). \quad (308)$$

Now we show that $G_2(x, y, z, t)$ is concave in x when $0 \leq z \leq t \leq \frac{1}{2}$, which generalizes the concavity of $G_1(x, y, z)$. To do this, we consider

$$\begin{aligned} & \frac{\partial^2}{\partial x^2} G_2(x, y, z, t) \\ &= -\frac{(yz - \bar{y}\bar{z})^2}{xyz + \bar{x}\bar{y}\bar{z}} - \frac{(\bar{y}z - y\bar{z})^2}{x\bar{y}z + \bar{x}y\bar{z}} - \frac{(y\bar{z} - \bar{y}z)^2}{xy\bar{z} + \bar{x}yz} - \frac{(\bar{y}\bar{z} - yz)^2}{x\bar{y}\bar{z} + \bar{x}yz} + \frac{(yt - \bar{y}\bar{t})^2}{xyt + \bar{x}\bar{y}\bar{t}} + \frac{(\bar{y}t - y\bar{t})^2}{x\bar{y}t + \bar{x}y\bar{t}} + \frac{(y\bar{t} - \bar{y}t)^2}{xy\bar{t} + \bar{x}yt} + \frac{(\bar{y}\bar{t} - yt)^2}{x\bar{y}\bar{t} + \bar{x}yt}, \end{aligned} \quad (309)$$

and

$$\frac{\partial}{\partial t} \left(\frac{\partial^2}{\partial x^2} G_2(x, y, z, t) \right) = \frac{\partial}{\partial t} \left(\frac{(yt - \bar{y}\bar{t})^2}{xyt + \bar{x}\bar{y}\bar{t}} + \frac{(y\bar{t} - \bar{y}t)^2}{x\bar{y}t + \bar{x}y\bar{t}} \right) + \frac{\partial}{\partial t} \left(\frac{(\bar{y}t - y\bar{t})^2}{x\bar{y}t + \bar{x}y\bar{t}} + \frac{(y\bar{t} - \bar{y}t)^2}{xy\bar{t} + \bar{x}yt} \right) \quad (310)$$

$$= \frac{-y^2 \cdot \bar{y}^2 \cdot (xy + \bar{x}\bar{y}) \cdot (1 - 2t)}{(xyt + \bar{x}\bar{y}\bar{t})^2 (x\bar{y}\bar{t} + \bar{x}yt)^2} + \frac{-y^2 \cdot \bar{y}^2 \cdot (x\bar{y} + \bar{x}y) \cdot (1 - 2t)}{(x\bar{y}t + \bar{x}y\bar{t})^2 (xy\bar{t} + \bar{x}yt)^2}. \quad (311)$$

Hence for $0 \leq t \leq \frac{1}{2}$,

$$\frac{\partial}{\partial t} \left(\frac{\partial^2}{\partial x^2} G_2(x, y, z, t) \right) \leq 0, \quad (312)$$

i.e., $\frac{\partial^2}{\partial x^2} G_2(x, y, z, t)$ is decreasing in t . Then we have for $0 \leq z \leq t \leq \frac{1}{2}$,

$$\frac{\partial^2}{\partial x^2} G_2(x, y, z, t) \leq \frac{\partial^2}{\partial x^2} G_2(x, y, z, z) = 0. \quad (313)$$

It implies $G_2(x, y, z, t)$ is concave in x when $0 \leq z \leq t \leq \frac{1}{2}$.

Combining (302) and (306), and utilizing the concavity of $G_2(x, y, z, t)$, we have

$$I(V_k; U_k | U_{k-1} Z_k) \geq \mu_k \left(H_2(\beta_k \star \tau_k) - H_2(\beta_k \star \tau_{k-1}) - G_2 \left(\sum_{v_k \in \mathcal{A}_k} \lambda_{v_k} d_{v_k}, \beta_k, \tau_k, \tau_{k-1} \right) \right) \quad (314)$$

$$= \mu_k (H_2(\beta_k \star \tau_k) - H_2(\beta_k \star \tau_{k-1}) - G_2(\alpha_k, \beta_k, \tau_k, \tau_{k-1})) \quad (315)$$

where α_k is given by (285) and satisfies (286) and (287). Therefore,

$$I(V_k; U_k | U_{k-1} Z_k) \geq \frac{\beta_k - D_k}{\beta_k - \alpha_k} (H_2(\beta_k \star \tau_k) - H_2(\beta_k \star \tau_{k-1}) - G_2(\alpha_k, \beta_k, \tau_k, \tau_{k-1})) \quad (316)$$

$$= \frac{\beta_k - D_k}{\beta_k - \alpha_k} (H_2(\beta_k \star \tau_k) - H_2(\beta_k \star \tau_{k-1}) - (H_4(\alpha_k, \beta_k, \tau_k) - H_4(\alpha_k, \beta_k, \tau_{k-1}))), \quad (317)$$

i.e., (290) holds.

Combining (245), (259) and (290) gives Theorem 9.

REFERENCES

- [1] L. Yu, H. Li, and W. Li, "Distortion bounds for source broadcast over degraded channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Barcelona, Spain, Jul. 2016, pp. 1834–1838.
- [2] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 623–656, Oct. 1948.
- [3] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 4, pp. 558–567, Oct. 1965.
- [4] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.
- [5] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.
- [6] S. Shamai, S. Verdú, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 564–579, Mar. 1998.
- [7] V. Prabhakaran, R. Puri, and K. Ramchandran, "Hybrid digital-analog codes for source-channel broadcast of Gaussian sources over Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4573–4588, Aug. 2011.
- [8] Z. Reznicek, M. Feder, and R. Zamir, "Distortion bounds for broadcasting with bandwidth expansion," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3778–3788, Aug. 2006.
- [9] C. Tian, S. Diggavi, and S. Shamai, "Approximate characterizations for the Gaussian source broadcasting distortion region," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 124–136, Jan. 2011.
- [10] L. Yu, H. Li, and W. Li, "Comments on 'Approximate characterizations for the Gaussian source broadcast distortion region'," To be published in *IEEE Trans. Inf. Theory*, [Online]. Available: <https://arxiv.org/abs/1607.08292>.
- [11] K. Khezeli, and J. Chen "A source-channel separation theorem with application to the source broadcast problem," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1764–1781, 2016.
- [12] K. Khezeli, and J. Chen "Outer bounds on the admissible source region for broadcast channels with correlated sources," *IEEE Trans. Inf. Theory*, vol. 61, pp. 4616–4629, Sep. 2015.
- [13] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 197–207, 1973.
- [14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.
- [15] P. Minero, S. H. Lim, and Y. H. Kim, "A unified approach to hybrid coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1509–1523, 2015.
- [16] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.
- [17] J. Nayak, E. Tuncel, and D. Gündüz, "Wyner-Ziv coding over broadcast channels: Digital schemes," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1782–1799, Apr. 2010.
- [18] Y. Gao and E. Tuncel, "Wyner-Ziv coding over broadcast channels: Hybrid digital/analog schemes," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5660–5672, Sept. 2011.
- [19] O. Shayevitz, and M. Wigger, "On the capacity of the discrete memoryless broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5660–5672, Sept. 2011.
- [20] S.-H. Lee and S.-Y. Chung, "A unified approach for network information theory," [Online]. Available: <http://arxiv.org/abs/1401.6023>.
- [21] M. H. Yassaee, M. R. Aref, and A. Gohari, "Achievability proof via output statistics of random binning," *IEEE Trans. Inf. Theory*, vol. 60, pp. 6760–6786, Nov. 2014.

- [22] P. Gács and J. Körner, "Common information is far less than mutual information" *Probl. Contr. Inform. Theory* vol. 2, no. 2, pp. 149-162, 1973.
- [23] R. Ahlswede and J. Körner, "On common information and related characteristics of correlated information sources," in *Proc. 7th Prague Conf. Inf. Theory*, 1974.